



Kauno technologijos universitetas
Matematikos ir gamtos mokslų fakultetas

**Sprendimų paramos modelių bei interaktyvios sistemos
kūrimas, vertinant pacientų, kuriems diagnozuotas priešinės
liaukos vėžys, mirtingumo riziką**

Baigiamasis magistro studijų projektas

Indrė Juškaitė

Projekto autorė

Doc. dr. Tomas Ruzgas

Vadovas

Kaunas, 2021



Kauno technologijos universitetas

Matematikos ir gamtos mokslų fakultetas

**Sprendimų paramos modelių bei interaktyvios sistemos
kūrimas, vertinant pacientų, kuriems diagnozuotas priešinės
liaukos vėžys, mirtingumo riziką**

Baigiamasis magistro studijų projektas

Taikomoji matematika (6211AX006)

Indrė Juškaitė

Projekto autorė

Doc. dr. Tomas Ruzgas

Vadovas

Doc. dr. Vytautas Janilionis

Recenzentas

Kaunas, 2021



Kauno technologijos universitetas

Matematikos ir gamtos mokslų fakultetas

Indrė Juškaitė

Sprendimų paramos modelių bei interaktyvios sistemos kūrimas, vertinant pacientų, kuriems diagnozuotas priešinės liaukos vėžys, mirtingumo riziką

Akademinio sąžiningumo deklaracija

Patvirtinu, kad:

1. baigiamąjį projektą parengiau savarankiškai ir sąžiningai, nepažeisdama(s) kitų asmenų autoriaus ar kitų teisių, laikydamasi(s) Lietuvos Respublikos autorių teisių ir gretutinių teisių įstatymo nuostatų, Kauno technologijos universiteto (toliau – Universitetas) intelektinės nuosavybės valdymo ir perdavimo nuostatų bei Universiteto akademinės etikos kodekse nustatytų etikos reikalavimų;
2. baigiamajame projekte visi pateikti duomenys ir tyrimų rezultatai yra teisingi ir gauti teisėtai, nei viena šio projekto dalis nėra plagijuota nuo jokių spausdintinių ar elektroninių šaltinių, visos baigiamojo projekto tekste pateiktos citatos ir nuorodos yra nurodytos literatūros sąrašė;
3. įstatymų nenumatytų piniginių sumų už baigiamąjį projektą ar jo dalis niekam nesu mokėjęs (-usi);
4. suprantu, kad išaiškėjus nesąžiningumo ar kitų asmenų teisių pažeidimo faktui, man bus taikomos akademinės nuobaudos pagal Universitete galiojančią tvarką ir būsiu pašalinta(s) iš Universiteto, o baigiamasis projektas gali būti pateiktas Akademinės etikos ir procedūrų kontrolieriaus tarnybai nagrinėjant galimą akademinės etikos pažeidimą.

Indrė Juškaitė

Patvirtinta elektroniniu būdu

Juškaitė Indrė. Sprendimų paramos modelių bei interaktyvios sistemos kūrimas, vertinant pacientų, kuriems diagnozuotas priešinės liaukos vėžys, mirtingumo riziką. Magistro studijų baigiamasis projektas / vadovas doc. dr. Tomas Ruzgas; Kauno technologijos universitetas, Matematikos ir gamtos mokslų fakultetas.

Studijų kryptis ir sritis (studijų krypties grupė): Taikomoji matematika (Matematikos mokslai).

Reikšminiai žodžiai: priešinė liauka, prostata, vėžys, priešinės liaukos vėžys, prostatos specifinis antigenas, limfmazgiai, internetinė svetainė, aprašomoji statistika, koreliacinė analizė, mašininis mokymasis, dirbtiniai neuroniniai tinklai.

Kaunas, 2021. 69 p.

Santrauka

Šiais laikais medicina tapo ypač svarbi visuomenei. Žmonės nori būti sveiki ir, sulaukę garbaus amžiaus, mirti dėl natūralių priežasčių, o ne paveikti sunkių ar mirtinų ligų. Tačiau Lietuvos Nacionalinio vėžio instituto duomenys atskleidė, kad Lietuvoje prostatos vėžys yra dažniausia vyrų onkologinė liga. Kas metai Lietuvoje nustatoma vis daugiau naujų priešinės liaukos vėžio atvejų. 2005 m. gruodžio 14 dieną patvirtintame Lietuvos Respublikos Sveikatos apsaugos ministro įsakyme dėl priešinės liaukos vėžio ankstyvoji diagnostikos finansavimo programos nustatyta, kad priešinės liaukos vėžio ankstyvoji diagnostika ir prostatos specifinio antigeno nustatymo paslauga bus teikiama vyrams nuo 50 iki 75 metų ir vyrams nuo 45 metų, jei jų tėvai ar broliai sirgo priešinės liaukos vėžiu. Taigi, vis daugiau pastangų dedama kuo anksčiau diagnozuoti šią ligą, o didėjant žmonių skaičiui pasaulyje to aktualumas dar labiau išauga.

Magistro baigiamajame projekte yra tiriamas ir analizuojamas priešinės liaukos vėžio mirtingumo rizikos vertinimas. Ši tema yra ypatingai aktuali gydytojams urologams, vertinantiems paciento, turinčio priešinės liaukos vėžio diagnozę, galimybes išgyventi trukmę. Tai padėtų gydytojui parinkti tinkamiausią paciento gydymo planą.

Tyrimo objektas – priešinės liaukos vėžio mirtingumo rizikos vertinimas.

Atsižvelgiant į tyrimo objektą, iškeltas pagrindinis darbo tikslas – sukurti automatizuotą ir statistiniais metodais paremtą modelį, skirtą vertinti pacientų, kuriems diagnozuotas priešinės liaukos vėžys, mirtingumo riziką.

Šiam tikslui pasiekti iškelti tyrimo uždaviniai:

1. Išanalizuoti pasaulyje atliktus tyrimus tyrimus ta pačia kaip magistro baigiamojo projekto tema.
2. Paruošti magistro baigiamojo projekto duomenų rinkinį analizei.
3. Sukurti magistro baigiamojo projekto tyrimo modelį.
4. Panaudojant sudarytą modelį ir programines priemones, sukurti interaktyvią sistemą skirtą priešinės liauko vėžio mirtingumo vertinimui.

Pirmoje šio projekto dalyje išanalizuoti pasaulyje atlikti tyrimai, susiję su šia tema. Antroje dalyje pateikta metodinė informacija apie duomenų ruošimą, koreliacinę analizę, mašininio mokymosi bei dirbtinių neuroninių tinklų metodus. Trečioje projekto dalyje, naudojant programinę įrangą, pagal sukurtus modelius išanalizuojami duomenys ir pateikiamos išvalgos. Taip pat aprašomas sukurtas internetinis puslapis ir automatizuotais bei statistiniais metodais paremta programa. Projekto duomenų analizė parodė, kad priešinės liaukos vėžys kaip mirties priežastis, nepriklausomai nuo amžiaus grupės, išlieka stabili. Sukurtas internetinis puslapis lengvai prieinamas adresu: <https://indremarijajuskait.wixsite.com/plv-modelis>. Projektui sukurta programa randama nurodytame internetiniame puslapyje.

Juškaitė Indrė. Development of decision support models and an interactive system for assessing the risk of mortality in patients diagnosed with prostate cancer. / supervisor doc. dr. Tomas Ruzgas; Faculty of Mathematics and Natural Sciences, Kaunas University of Technology.

Study field and area (study field group): Applied Mathematics (Mathematical Sciences).

Keywords: prostate, cancer, prostate cancer, prostate-specific antigen, lymph nodes, website, descriptive statistics, correlation analysis, machine learning, artificial neural networks.

Kaunas, 2021. 69.

Summary

Nowadays, medicine plays a particularly important role in our society. The aim is to pass away in peace due to the natural causes rather than a serious or deadly disease. However, data from the Lithuania's National Cancer Institute revealed that the most common oncological disease amongst the men in this country is prostate cancer. Each year, there is an increase in the number of cases of patients with prostate cancer. On the 14th of December 2005, Lithuania's Minister of Health released an order about the early diagnosis of prostate cancer financing program, which stated that the early diagnosis of prostate cancer and prostate-specific antigen detection service will be provided to men aged 50 to 75 years old, and men aged 45 years old if their parents or brothers had a history of prostate cancer. Consequently, due to the rising numbers in the population, there is an increase in the relevance and therefore the effort that is being put into the diagnosis of this disease.

This project focuses on the topic of the risk assessment of prostate cancer mortality. The topic is particularly relevant for urologists assessing the survival period of a patient diagnosed with prostate cancer. It would help the doctors choose the most appropriate treatment plan for the patient.

The main problem of the study is being expressed in the question: How to quickly and effectively determine the life expectancy in patients with prostate cancer?

The objective of the study is the risk assessment of prostate cancer mortality.

Considering the subject of the study, the main goal is to develop an automated and statistic-based model to assess the risk of mortality in patients diagnosed with prostate cancer.

For achieving this goal, the following research tasks have been set:

1. To analyse researches that were carried out on the topic of the master's final project around the world.
2. To prepare the data set for analysis.
3. To create a research model.
4. To create an interactive system for the estimation of prostate cancer mortality using the developed model and software tools.

The first part of the project critically examines the researches that were conducted in relation to this topic. The second part provides methodological information on data preparation, correlation analysis,

machine learning, and artificial neural network. In the third part of the final thesis, using the software, based on the created models, the data is analysed, and insights are presented according to the developed models. The created website is being outlined, and the program based on automated and statistical methods is also being described. Data analyses showed that prostate cancer, as a cause of death, remains stable regardless of age group. The created website is easily accessible at: <https://indremarijajuskait.wixsite.com/plv-modelis>. The program created for this project is found on this website.

Turinys

Lentelių sąrašas	9
Paveikslų sąrašas	10
Santrumpų ir terminų sąrašas	11
Įvadas.....	12
1. Literatūros apžvalga	13
1.1. Vėžys	13
1.2. Priešinės liaukos vėžys	16
1.3. Internetinių svetainių kūrimas, WIX	20
1.4. Statistinės analizės sistema, R	21
2. Projekto tyrimo metodai.....	23
2.1. Aprašomoji statistika	23
2.2. Koreliacinė analizė	25
2.3. Mašininis mokymasis ir dirbtiniai neuroniniai tinklai.....	27
2.4. Daugialypė tiesinė regresija	32
3. Projekto tyrimo rezultatai ir aptarimas	33
3.1. Duomenų rinkinio parengimas	33
3.2. Statistinės analizės modelis	34
3.3. Koreliacinės analizės modelis	38
3.4. Internetinis puslapis.....	40
3.5. Automatizuotas ir statistiniais metodais paremtas modelis, programa	42
Išvados ir rekomendacijos	51
Literatūros sąrašas	52
Priedai.....	55
1 priedas. Diagnozuotų onkologinių susirgimų skaičius ir sergamumo rodikliai Lietuvoje 2015 metais. Vyrai.....	55
2 priedas. Pagrindiniai vėžio tipai pagal naujus atvejus (sergamumą) ir mirtis (mirtingumą) pagal lytį kiekviename iš 20 pasaulio regionų 2020 m. [6]	57
3 priedas Aprašomosios statistikos rezultatų vizualizacija R programinėje įrangoje	59
4 priedas Skirtingų amžiaus grupių pacientų mirties rodiklis, vienetais	60
5 priedas Neišgyvenusių skirtingų amžiaus grupių pacientų mirties priežastis, vienetais	61
6 priedas Internetinio puslapio „PLV“ skirtuko vaizdas	62
7 priedas priedas Internetinio puslapio „Diagnostika“ skirtuko vaizdas 1	63
8 priedas Internetinio puslapio „Diagnostika“ skirtuko vaizdas 2	64
9 priedas Naudojimosi instrukcijos skirtuko vaizdas 1	65
10 priedas Naudojimosi instrukcijos skirtuko vaizdas 2	66
11 priedas Statistinės ir koreliacinės analizės programos kodai	67

Lentelių sąrašas

1 lentelė. Pasaulio gyventojų prognozė (2020–2050 m.) [7].....	15
2 lentelė. Pradinių duomenų atributai ir jų reikšmės	33
3 lentelė. Duomenų failo tuščių atributų skaičius.....	34
4 lentelė. Aprašomosios statistikos rezultatai, skaitinio tipo kintamieji.....	34
5 lentelė. Aprašomosios statistikos rezultatai, ranginio tipo kintamieji.....	35
6 lentelė. Spirmeno koreliacijos reikšmės	39
7 lentelė. Interaktyvios sistemos funkciniai reikalavimai	40
8 lentelė. Testavimo duomenų „SURVIVAL“ atributas	46
9 lentelė. 1 metų modelio pasirinkimas	46
10 lentelė. 2 metų modelio pasirinkimas	47
11 lentelė. 5 metų modelio pasirinkimas	47

Paveikslų sąrašas

1 pav. Normalios ląstelės.....	13
2 pav. Vėžinės ląstelės.....	14
3 pav. Normalių ir nenormalių ląstelių augimas [4].....	14
4 pav. Prostatos vėžys.....	16
5 pav. DNR dviguba spirale [10].....	17
6 pav. Priešinės liaukos vėžio stadijos [14].....	20
7 pav. Programinės įrangos galimybių ir mokymosi kreivė [20]	22
8 pav. Skaitinės charakteristikos [21].....	23
9 pav. Koreliacijos pavyzdys [25]	25
10 pav. Schema, rekomenduojanti, kuri koreliacijos koeficientą pasirinkti [27]	27
11 pav. Klasikinio programavimo realizacija.....	28
12 pav. Mašininio mokymosi realizacija.....	28
13 pav. Neuroninio tinklo mokymo ir testavimo fazės [29].....	29
14 pav. Biologinio neurono sandara [31]	30
15 pav. Dirbtinio neurono sandara [31].....	30
16 pav. Slenkstinė funkcija [32].....	31
17 pav. Programinės įrangos R, funkcijos „str“ rezultatai apie atributus.....	34
18 pav. Procentalus pacientų pasiskirstymas pagal amžiaus grupes	35
19 pav. Pacientų PSA rodiklio vidurkis amžiaus grupėse	36
20 pav. Pacientų mirties rodiklis amžiaus grupėse, procentais	37
21 pav. Neišgyvenusių skirtingo amžiaus pacientų mirties priežastis, procentais	37
22 pav. Amžiaus pasiskirstymo grafikas	38
23 pav. PSA pasiskirstymo grafikas	38
24 pav. Išgyvenamumo pasiskirstymo grafikas.....	39
25 pav. Spirimo koreliacija trims kintamiesiems.....	39
26 pav. Nuoroda į Nacionalinio vėžio instituto internetinį puslapį.....	40
27 pav. Skirtukas „Diagnostika“	40
28 pav. Skirtukas „Naudojimosi instrukcija“	41
29 pav. Skirtukas „Prognozavimo įrankis“	41
30 pav. Skirtukas „Kontaktai“	42
31 pav. Projekto interaktyvios sistemos funkcionalumas.....	42
32 pav. Programos „PLV modelis“ vartotojo sąsajos vaizdas	43
33 pav. Programos „PLV modelis“ išvešėjusių limfmazgių skaičiaus pasirinkimai.....	43
34 pav. Programos „PLV modelis“ patloginės vėžio stadijos pasirinkimai.....	44
35 pav. Programos „PLV modelis“ limfmazgių būklės pasirinkimai	44
36 pav. Programos „PLV modelis“ pašalintos prostatos būklės pasirinkimai	45
37 pav. Aprašomosios statistikos rezultatų vizualizacija R programinėje įrangoje testavimo duomenims	45
38 pav. Programos įvesčių ir rezultatų vaizdas	48
39 pav. Neuroninis tinklas 1 metams	49
40 pav. Neuroninis tinklas 2 metams	49
41 pav. Neuroninis tinklas 5 metams	50

Santrumpų ir terminų sąrašas

Santrumpos:

Doc. – docentas;

Lekt. – lektorius;

Prof. – profesorius.

PAS – prostatos specifinis antigenas.

Terminai:

Onkogenai – genai esantys žmogaus organizme [1].

Naviko slopinimo genai – baltymus koduojantys genai [1].

Vėžys – tai liga, kai nenormalios ląstelės dalijasi nekontroliuojamos ir gali įsiskverbti į netoliese esančius audinius [2].

Metastazė – vėžinių ląstelių plitimas į tolesnes organizmo vietas [2].

Navikas – neribotai dalijančiasi ir pakitusių ląstelių darinys [2].

Prostatos specifinis antigenas – prostatos gaminama baltyminė medžiaga.

DNR –yra viena iš dviejų molekulių rūšių, koduojančių genetinę informaciją [9].

GNU – nemokamos programinės įrangos rinkinys. Jis naudojamas kaip operacinė sistema arba kartu su kitomis operacinėmis sistemomis.

Aprašomoji statistika – duomenų sisteminimo ir grafinio vaizdavimo metodai [21].

Koreliacija – žymi dviejų ir daugiau kiekybinių kintamųjų ryšius [25].

Mašininis mokymasis – kompiuterių mokslo sritis kuri naudojami duomenimis ir išmoksta atpažinti specifines duomenų grupes ar kita.

Dirbtinis neuroninis tinklas – biologiniais neuroniniais tinklais paremtas mašininio mokymosi algoritmas [28].

Regresija – duomenų gavimo technika.

Įvadas

Vėžys yra plačiai paplitusi ir puikiai žinoma liga ne tik Lietuvoje, bet ir visame pasaulyje. Vienas iš dažniausiai pasitaikančių vėžio, kuriuo serga vyrai, tipų yra priešinės liaukos vėžys. Deja, mokslininkai dar nėra gerai ištyrę šios ligos ir jos gydymas nėra šimtu procentų veiksmingas visais atvejais. Kuo toliau, tuo labiau augant žmonių skaičiui pasaulyje, gydytojams ir mokslininkams svarbu suprasti šios ligos priežastis, taip pat vyrams, turintiems priešinės liaukos vėžio diagnozę, aproksimuoti laikotarpį, kurį jie galėtų išgyventi. Su šiuo iššūkiu susiduria ir Lietuvos gydytojai urologai.

Pagrindinė tyrimo problema išreiškiama klausimu: kaip galima greitai ir efektyviai aproksimuoti laikotarpį, kurį išgyvens priešinės liaukos vėžiu sergantis vyras?

Siekiant kuo tiksliau diagnozuoti pacientus, šiam magistro baigiamajam projektui buvo iškeltas tikslas.

Tyrimo tikslas – sukurti automatizuotą ir statistiniais metodais paremtą modelį, skirtą vertinti pacientų, kuriems diagnozuotas priešinės liaukos vėžys, mirtingumo riziką.

Siekiant tikslo, sprendžiami uždaviniai:

1. Išanalizuoti pasaulyje atliktus tyrimus ta pačia kaip magistro baigiamojo projekto tema.
2. Paruošti magistro baigiamojo projekto duomenų rinkinį analizei.
3. Sukurti magistro baigiamojo projekto tyrimo modelį.
4. Panaudojant sudarytą modelį ir programines priemones, sukurti interaktyvią sistemą skirtą priešinės liauko vėžio mirtingumo vertinimui.

Magistro baigiamajame darbe naudojami metodai:

- mokslinės literatūros analizė;
- bendroji duomenų analizė;
- koreliacinė analizė;
- dirbtiniai neuroniniai tinklai.

Projektą sudaro teorinė ir praktinė dalys. Teorinėje dalyje išanalizuoti matematiniai metodai, programinė įranga ir atlikti tyrimai darbo tema. Praktinėje dalyje analizuojami realūs duomenys. Sukurtas lengvai prieinamas ir informatyvus internetinis puslapis ir automatizuotas statistiniais metodais paremtas modelis.

1. Literatūros apžvalga

1.1. Vėžys

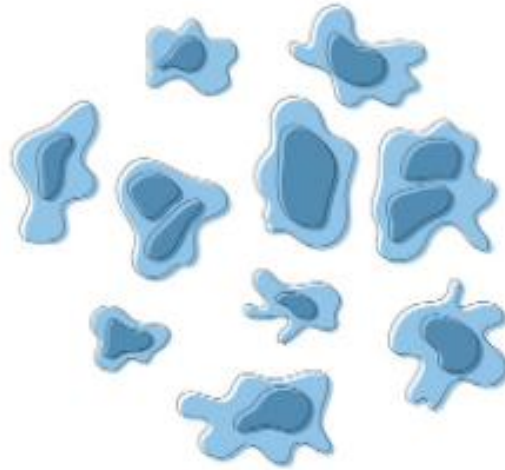
Per visą pasaulio istoriją ne tik žmogus, bet ir kiti gyvūnai sirgo vėžiu. Pirmasis vėžio ligos apibrėžimas buvo sugalvotas graikų gydytojo Hipokrato dar 460–370 m. pr. Kr. [1]. Šiuo metu Nacionalinis vėžio institutas vėžį apibrėžia kaip ligos terminą, kurio metu nenormalios ląstelės dalijasi nekontroliuodamos ir gali išskverbtis į netoliese esančius audinius [2]. Per kraują ar limfą vėžinės ląstelės gali išplisti ir į kitus organus. Yra keletas pagrindinių vėžio tipų [2]:

- Karcinoma yra vėžys, prasidedantis epitelio ląstelėse (odoje arba audiniuose, kurie riboja arba dengia vidaus organus).
- Sarkoma yra vėžys, prasidedantis kauluose, kremzlėse, riebaluose, raumenyse, kraujagyslėse ar kituose jungiamuosiuose audiniuose.
- Leukemija yra vėžys, kuris prasideda kraują formuojančiuose audiniuose, pavyzdžiui, kaulų čiulpuose, ir sukelia nenormaliai didelį kraujo ląstelių dalijimąsi.
- Limfoma yra vėžys, prasidedantis limfinio audinio ląstelėse.
- Centrinės nervų sistemos vėžys, prasideda smegenų ir nugaros smegenų audiniuose.

Kiekvieno žmogaus organizmą sudaro trilijonai ląstelių. Šios ląstelės per visą žmogaus gyvenimą įprastai auga ir dalijasi pagal poreikį. Kai ląstelės yra nenormalios arba sensta, jos žūva. Vėžys prasideda tada, kai šis procesas sutrinka ir ląstelės nuolat kuria naujas ląsteles. Tuo tarpu senos ar nenormalios ląstelės nemiršta tada, kai jos turėtų. Kai vėžinės ląstelės auga nekontroliuojamos, jos gali išstumti normalias ląsteles [3] (žr. 1 pav. ir 2 pav.).

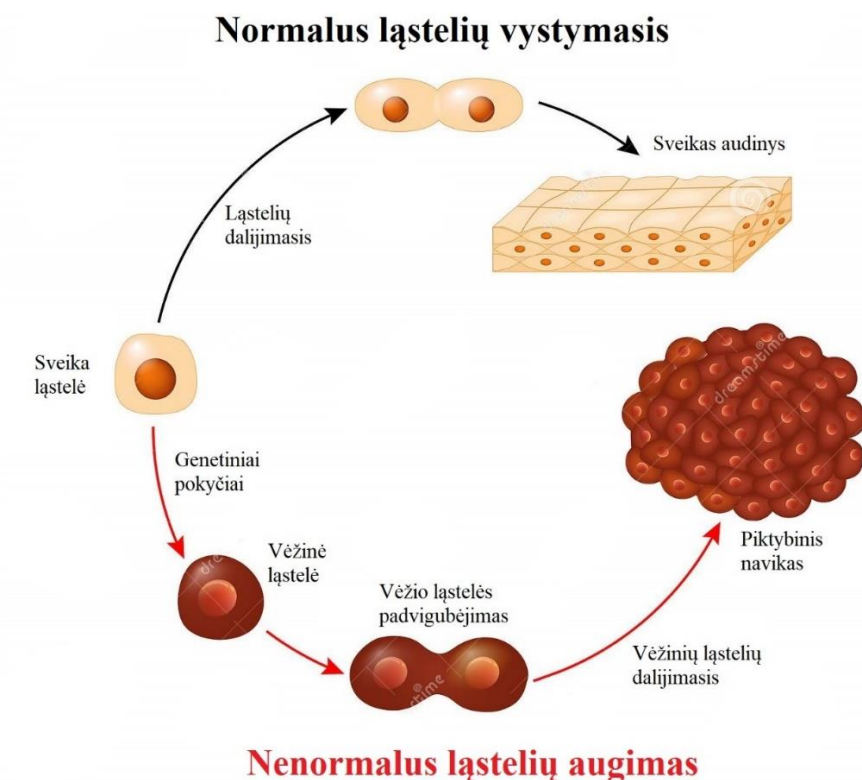


1 pav. Normalios ląstelės



2 pav. Vėžinės ląstelės

Įprastai organizme ląstelės dalijasi normaliai, taigi iš vienos sveikos ląstelės atsiranda dvi sveikos ląstelės. Šis procesas kartojasi tol, kol sukuriamas sveikas audinys organizme. Taip pat normalioms ląstelėms pasenus jos žūsta, o audinys išlieka sveikas. Kitu atveju, jeigu ląstelėje įvyksta tam tikrų genetinių pokyčių, ji tampa vėžine ląstele. Vėžinė ląstelė taip pat dalijasi ir tokiu būdu organizme susiformuoja piktybinis navikas. (žr. 3 pav.).



3 pav. Normalių ir nenormalių ląstelių augimas [4]

Vėžys yra viena iš dažniausių bei sunkiausių ligų ne tik Lietuvoje, bet ir pasaulyje. Amerikos vėžio draugijos teigimu, vienas iš trijų asmenų Jungtinėse Amerikos Valstijose serga šia liga [1]. Nacionalinio vėžio instituto paskutiniais duomenimis, diagnozuotų naujų onkologinių susirgimų skaičius ir sergamumo rodiklis Lietuvoje 2015 metais siekė 37089 atvejus [5]. Iš jų vyrų 19004, o

moterų 18085 [5]. Lietuvoje viena iš dažniausiai pasitaikančių vėžio rūšių, diagnozuojamų vyrams, yra priešinės liaukos vėžys. 2015 metų duomenimis buvo aptikti 2967 nauji priešinės liaukos vėžio atvejai [5] (žr. 1 priedą).

Per 2020 metus iš 185 pasaulio šalių buvo surinkti duomenys apie vėžio paplitimą ir mirtingumą pagal amžiaus grupes. Apskaičiuota, kad 2020 metais visame pasaulyje buvo 19,3 mln. naujų vėžio atvejų ir beveik 10 mln. mirčių nuo šios ligos. Tuo metu dažniausiai visame pasaulyje diagnozuotas vėžys buvo moterų krūties vėžys – net 2,26 mln. atvejų, plaučių vėžys – 2,21 mln. atvejų ir prostatos vėžys – 1,41 mln. atvejų. Be to, apskaičiuota, kad didžiausią mirtingumą sukėlė plaučių vėžys – 1,79 mln. mirčių, kepenų vėžys – 830 tūkst. mirčių ir skrandžio vėžys – 769 tūkst. mirčių. Iki 75 metų amžiaus rizika susirgti vėžiu yra 20 %, o rizika mirti nuo šios ligos yra 10 %. Atlikto tyrimo metu buvo nustatyta, jog kas penktas asmuo pasaulyje susirgs viena iš vėžio formų, o kas dešimtas mirs nuo šios ligos [6]. Pažvelgus į šiuos rodiklius yra akivaizdu, kad vėžys yra viena iš aktualiausių šių dienų medicinos problemų. Tai lemia ne tik milžiniški susirgimų skaičiai, bet ir aukštas mirtingumo lygis.

2020 metais pasaulyje buvo apie 7,8 milijardai gyventojų, prognozuojama, kad 2025 metais pasaulyje bus apie 8,2 milijardai žmonių, o 2050 – apie 9,7 milijardai gyventojų (žr. 1 lentelę) [7]. Akivaizdu, kad kiekvienais metais gyventojų skaičius pasaulyje vis labiau auga. Atsižvelgiant į šiuos rodiklius galima teigti, jog tikėtina, kad ir vėžio susirgimų skaičius bėgant laikui tik išaugs, nes šia liga galima susirgti ne tik dėl tam tikrų žalingų įpročių, bet ir dėl paveldimumo.

1 lentelė. Pasaulio gyventojų prognozė (2020–2050 m.) [7]

Metai	Populiacija	Metinis % pokytis	Metinis pokytis	Amži aus medi ana	Vaisingu mo rodiklis	Tank is (P/K m ²)	Miestų Pop. %	Miestų populiacija
2020	7794798739	1.10 %	83000320	31	2.47	52	56.2 %	4378993944
2025	8184437460	0.98 %	77927744	32	2.54	55	58.3 %	4774646303
2030	8548487400	0.87 %	72809988	33	2.62	57	60.4 %	5167257546
2035	8887524213	0.78 %	67807363	34	2.70	60	62.5 %	5555833477
2040	9198847240	0.69 %	62264605	35	2.77	62	64.6 %	5938249026
2045	9481803274	0.61 %	56591207	35	2.85	64	66.6 %	6312544819
2050	9735033990	0.53 %	50646143	36	2.95	65	68.6 %	6679756162

Kalbant apie Europos vėžio susirgimo rodiklius, pagal atvejų dažnumą pirmoje vietoje yra vyrų priešinės liaukos vėžys, antroje vietoje – plaučių vėžys, o trečioje – ne melanominis odos vėžys. Tačiau Europos vyrų mirtingumo nuo vėžio duomenimis pirmoje vietoje – plaučių vėžys, antroje vietoje – priešinės liaukos vėžys, o trečioje storosios žarnos vėžys (žr. 2 priedo 1, 2 lapus) [6]. Europos mastu pagal moterų susirgimų skaičių pirmoje vietoje yra krūties vėžys, antroje vietoje – plaučių vėžys, o trečioje vietoje – storosios žarnos vėžys. Be to, pagal mirtingumo duomenis Europoje pirmoje vietoje yra moterų krūties vėžys, antroje vietoje yra plaučių vėžys ir trečioje vietoje yra storosios žarnos vėžys (žr. 2 priedo 1, 2 lapus) [6]. Atsižvelgiant į šiuos duomenis matyti, jog vienas iš daugiausiai paplitusių vėžio tipų Europoje yra priešinės liaukos vėžys.

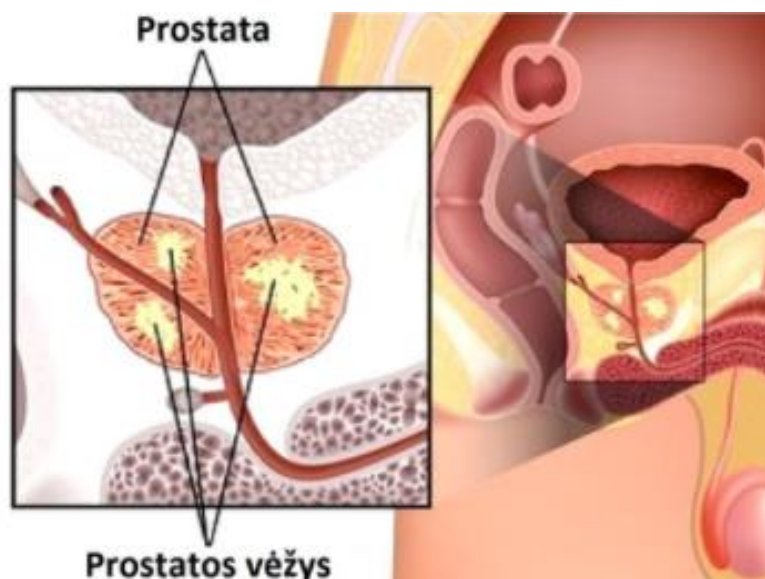
Akivaizdu, kad, prabėgus net tūkstančiams metų nuo šios ligos atradimo, nei mokslininkams, nei gydytojams dar nepavyko užkirsti kelio šios ligos plitimui. Šiuo metu pasaulyje dar nėra išrastos

vakcinos, kuri užtikrintų, kad asmuo nesusergs šia liga. Taip pat dar nėra šimtu procentų veiksmingo gydymo metodo, leidžiančio pacientams visai pasveikti ir užtikrinti, jog vėžys neatsinaujins.

1.2. Priešinės liaukos vėžys

Pagal Nacionalinio vėžio instituto duomenis, šiuo metu Lietuvoje priešinės liaukos vėžys yra dažniausia vyrų onkologinė liga [8]. Lietuvoje ši vyrų liga per metus nustatoma netgi apie 2000 kartų, o sergamumas ja nuolat auga [8]. Priešinės liaukos vėžys dažniausiai aptinkamas vyrams, vyresniems nei 50 metų amžiaus. Jaunesniems vyrams ši liga taip pat aptinkama, tačiau daug rečiau. Būtent priešinės liaukos vėžys yra labai klatinga diagnozė tuo atžvilgiu, kad ji, kitaip negu kitų organų vėžys, gali daugelį metų būti priešinėje liaukoje ir nesukelti jokių simptomų [8]. Dažnai pasitaiko, jog prostatoje navikai auga lėtai. Būtent dėl šios priežasties jaunesniems nei 50 metų amžiaus vyrams ši liga gali anksti nesukelti problemų ar netgi niekada taip ir nebūti diagnozuota. Žinoma, yra atvejų, kai ši vėžio forma progresuoja sparčiai ir kartais išplinta į kitus organus [8].

Prostata yra liauka, randama tik vyrų organizmuose. Vyro organizme prostata yra žemiau šlapimo pūslės (organas, kuriame laikomas šlapimas) ir prieš tiesiąją žarną (paskutinė žarnyno dalis) [1]. Už prostatos yra liaukos, vadinamos sėklinėmis pūslelėmis. Sėklinėse pūslelėse susidaro didžioji spermos skysčio dalis [1]. Šlaplė yra vamzdelis, kuris eina per prostatos centrą [1]. Prostatos vėžys prasideda tuomet, kai prostatos liaukos ląstelės pradeda nekontroliuojamai dalintis ir augti [1] (žr. 4 pav.). Vyriui senstant, prostatos dydis gali pasikeisti. Jaunesnių vyrų prostata yra maždaug graikinio riešuto dydžio, o vyresnių vyrų ji gali būti daug didesnė [1].



4 pav. Prostatos vėžys

Dažniausiai nustatoma adenokarcinomos tipo prostatos vėžys. Šis tipas išsivysto iš liaukos ląstelių, tai yra ląstelių, iš kurių susidaro prostatos skystis [1].

Kiti, rečiau pasitaikantys, prostatos vėžio tipai yra šie [1]:

- smulkiųjų ląstelių karcinomos;
- neuroendokrininiai navikai (išskyrus smulkiųjų ląstelių karcinomas);
- pereinamųjų ląstelių arba uroteliniai;
- sarkoma.

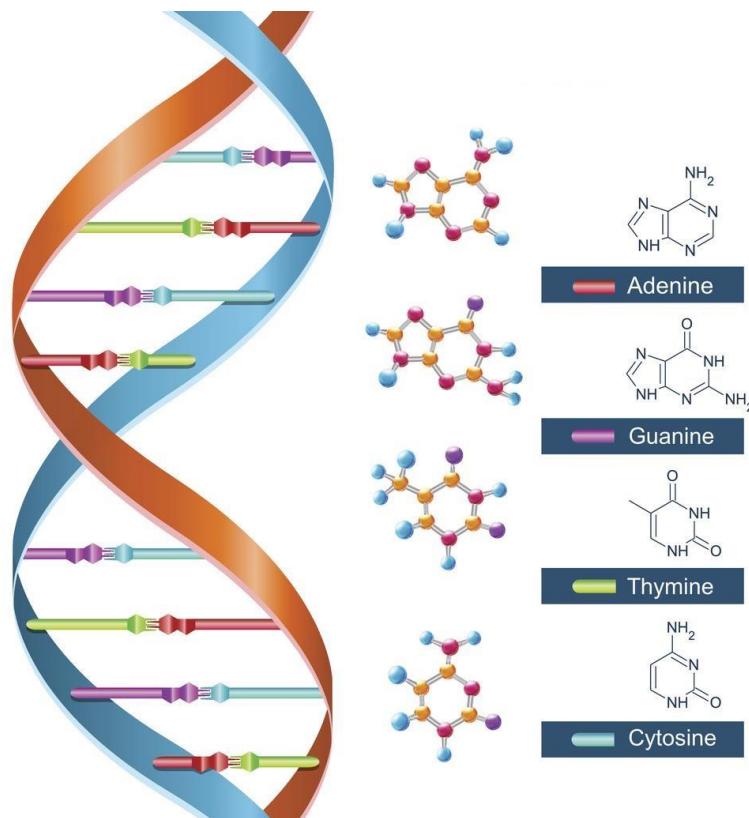
Tam tikri priešinės liaukos vėžio dariniai auga ir plinta sparčiai, tačiau dauguma auga pamažu. Iš tiesų, autopsijos tyrimai rodo, kad daugelis vyresnių vyrų (ir net kai kurių jaunesnių vyrų), mirusių dėl kitų priežasčių, taip pat sirgo prostatos vėžiu [1]. Nepaisant to, per visą tokių žmonių gyvenimą ši liga niekada jų taip ir nepaveikė.

Visos priešinės liaukos vėžio atsiradimo priežastys dar nėra aiškios, bet yra žinomi keli rizikos veiksniai, galintys lemti susirgimą [8]:

- Artimi giminės (tėvas, brolis, dėdė) serga ar sirgo priešinės liaukos vėžiu.
- Šeimos moterims buvo diagnozuotas krūties vėžys, ypačingai tuo atveju, jei jos susirgo jaunesnio negu 40 metų amžiaus. Šis faktorius gali reikšti, jog šeimos palikuonys paveldi mutavusį geną.
- Kai kurių etninių grupių vyrai yra didesnės susirgimo rizikos grupėje, pavyzdžiui, afroamerikiečiai.
- Riebus maistas ir nedidelis šviežių vaisių ir daržovių vartojimas.

Tačiau vieno ar kelių rizikos faktorių turėjimas dar nelemia, kad vyras susirgs šia liga. Daugelis žmonių, turinčių vieną ar net kelis rizikos veiksnius, gali taip niekada ir nesusirgti vėžiu, o kiti, susirgę vėžiu, galėjo turėti mažai žinomų rizikos veiksnių arba jų išvis neturėti.

Paprastai prostatos vėžį sukelia normalios prostatos ląstelės deoksiribonukleorūgšties DNR pokyčiai [1]. DNR yra viena iš dviejų molekulių rūšių, koduojančių genetinę informaciją [9]. DNR yra dvigubos grandinės molekulė, kurią tarpusavyje laiko silpni vandenilio ryšiai tarp bazinių nukleotidų porų [9]. Molekulė sudaro dvigubą spiralę, kurioje dvi DNR grandinės sukasi viena apie kitą (žr. 5 pav.) [9].



5 pav. DNR dviguba spiralė [10]

Taigi DNR yra cheminė medžiaga ląstelėse, sudaranti mūsų genus, o genai kontroliuoja ląstelių funkcionavimą [1]. Tam tikri genai kontroliuoja, kada ląstelės turi dalintis į naujas ląsteles, augti bei mirti [1]:

- Onkogenai, tai tam tikri genai, padedantys ląstelėms augti, dalytis ir išlikti gyvoms.
- Naviko slopinimo genai, kontroliuojantys ląstelių augimą, taisantys DNR klaidas arba priverčiantys ląsteles žūti tinkamu metu.

Vėžys gali būti sukeltas DNR mutacijos, kuri leidžia onkogenams tapti aktyviems arba sustabdo naviko slopinimo genų veiklą. Tokio tipo genų pokyčiai gali sukelti nekontroliuojamą ląstelių augimą. DNR pokyčiai gali būti paveldimi iš tėvų arba įgyjami per asmens gyvenimą. Kai kurios genų mutacijos yra perduodamos iš kartos į kartą, tiksliau, paveldimos, ir yra visose kūno ląstelėse. Mokslininkai mano, jog paveldimi genų pokyčiai turi įtakos atsirasti maždaug 10 % prostatos vėžio atvejų [1].

Per žmogaus gyvenimą kai kurie genai mutuoja ir palikuonims tokio tipo mutacijos neperduodamos. Šie pokyčiai būna tik ląstelėse, kurios yra kilusios iš pirminės mutavusios ląstelės. Tai vadinama įgytomis mutacijomis [1]. Pastebėta, kad dauguma genų mutacijų, susijusių su prostatos vėžiu, išsivysto vyro gyvenimo metu, o ne paveldima [1]. Kas kartą, kai ląstelė dalijasi į 2 naujas ląsteles, ji privalo nukopijuoti savo DNR. Toks procesas yra netobulas, taigi kartais gali pasitaikyti klaidų ir naujoje ląstelėje gali likti defektinė DNR. Kolkas dar neaišku, kaip dažnai šie DNR pokyčiai yra atsitiktiniai ir kada jiems įtaką daro kiti veiksniai. Apskritai, kuo greičiau prostatos ląstelės auga ir dalijasi, tuo daugiau galimybių atsirasti mutacijoms. Dėl šios priežasties viskas, kas pagreitina ląstelių dalijimosi procesą, gali padidinti prostatos vėžio atsiradimo riziką [1]. Pavyzdžiui, androgenai (vyriški hormonai), tokie kaip testosteronas, skatina prostatos ląstelių augimą [1]. Taigi yra nustatyta, kad didesnis androgenų kiekis vyrų organizme gali sukelti prostatos vėžį ar bent jau padidinti šios ligos susirgimo riziką. Kai kurie tyrimai parodė, kad vyrai, turintys aukštą kito hormono, į insuliną panašaus augimo faktoriaus kiekį, dažniau serga prostatos vėžiu [1]. Kiti tyrimai parodė, kad uždegimas prostatoje gali būti susijęs su šio organo vėžiu [1]. Viena iš teorijų yra ta, kad uždegimas gali sukelti ląstelių DNR pažeidimus, kurie gali padėti normaliai ląstelei tapti vėžine ląstele [1].

Pirminis prostatos vėžio nustatymas atliekamas tikrinant prostatos specifinio antigeno (PSA) serumo kiekį vyro organizme. Šio tyrimo rezultatai gaunami iš kraujo. PSA yra baltyminė medžiaga. Šią medžiagą gamina prostata. Lietuvos Respublikos sveikatos apsaugos ministro įsakymo 8.5. punkte buvo patvirtinta PSA norma Lietuvoje [11]. „8.5. skirti pacientui siuntimą konsultuotis pas urologą (atlikti priešinės liaukos biopsiją), jeigu PSA kiekis viršija 3 ng/ml.“ [11]. Taip pat, tame pačiame įsakyme 9 punkte buvo nustatyta, kad „informavimo apie priešinės liaukos vėžio ankstyvąją diagnostiką ir PSA nustatymo paslauga teikiama vyrams nuo 50 iki 75 metų ir vyrams nuo 45 metų, jei jų tėvai ar broliai sirgo priešinės liaukos vėžiu“ [11]. Dėl šių priežasčių įprastai kasmetiniai PSA tyrimai atliekami vyrams nurodyta tvarka.

Daugeliu atvejų pasireiškia simptominė lokaliai išplitusi prostatos vėžio liga ar tolimos metastazės, ypač vyrams, kuriems nebuvo reguliariai atliekami PSA tyrimai [12]. PSA tyrimas yra naudojamas kaip testavimo priemonė vyrams, neturintiems ligos simptomų. Tai, kad PSA tapo pirmuoju plačiai naudojamu prostatos vėžio patikros testu, pakeitė pastebėtą ligos dažnį ir pasiskirstymą [12]. Dešimtojo dešimtmečio pradžioje šalyse, kurios diagnozei naudojo PSA, ypač Jungtinės Valstijos, pastebėjo nuolatinį prostatos vėžio dažnio augimą [12]. Standartinis prostatos vėžio klasifikavimas atliekamas pagal „Gleason“ vertinimo sistemą, kurią pirmą kartą aprašė Gleasonas ir Mellingeris

1974 metais [12]. 2002 metais prostatos adenokarcinomos klasifikacija yra pagrįsta šiomis kategorijomis [12]:

- T (pirminio naviko);
- N (limfmazgių);
- M (metastazių).

T kategorija yra pagrįsta klinikiniais tyrimais, vaizdais, endoskopija, biopsija ir biocheminiais tyrimais, N kategorija pagrįsta klinikiniais tyrimais ar vaizdais, o M kategorija pagrįsta klinikiniais tyrimais, vaizdais, griaučių tyrimais ir biocheminiais tyrimais [12]. Skaičiai arba raidės po T, N ir M pateikia daugiau informacijos apie priešinės liaukos vėžio veiksmus. Kuo didesnis skaičius, tuo labiau pažengusi ir vėžio stadija. Gydytojams urologams nustatius T, N ir M kategorijas, ši informacija sujungiama su PSA rodikliu procese, kuris yra vadinamas etapų grupavimu, siekiant gauti bendrą vėžio stadiją [13]. Yra nustatytos keturios pagrindinės priešinės liaukos vėžio stadijos (žr. 6 pav.). Kuo žemesnė stadija yra nustatyta pacientui, tuo mažiau vėžys yra išplitęs.

I stadijos apibūdinimas (žr. 6 pav.) [13]:

- Vėžys nėra išplitęs į netoliese esančius limfmazgius ar kitur organizme.
- PSA lygis yra mažesnis nei 10.
- Navikas gali būti jaučiamas atliekant tiesiosios žarnos tyrimą arba matomas atliekant vaizdo tyrimą, pavyzdžiui, transrektalinį ultragarsą, ir yra tik vienoje prostatos pusėje (kairėje arba dešinėje).
- Prostata pašalinta operacijos būdu, o auglys buvo tik prostatoje.

II stadijos apibūdinimas (žr. 6 pav.) [13]:

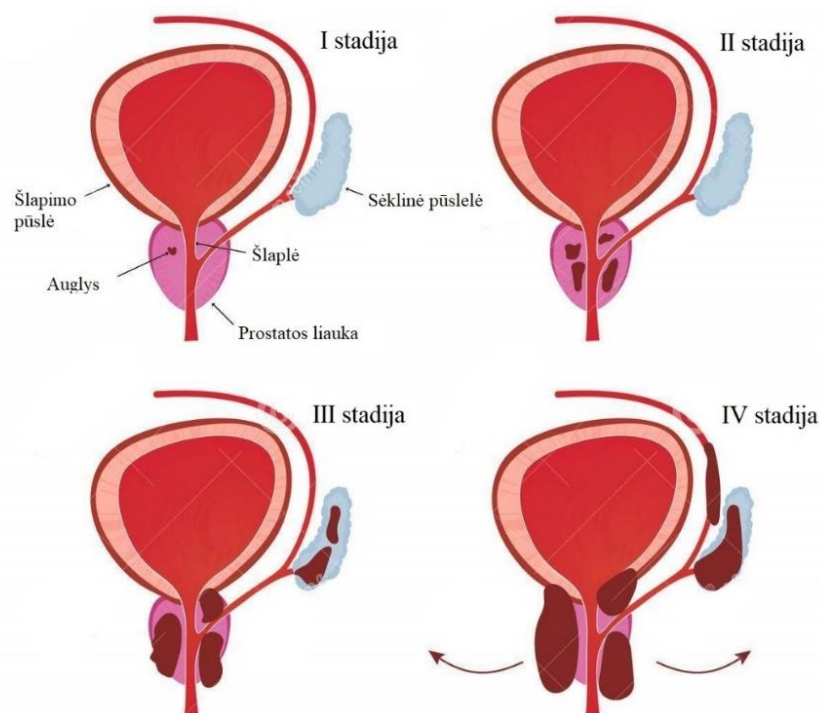
- Vėžys nėra išplitęs į netoliese esančius limfmazgius ar kitur organizme.
- PSA lygis yra tarp 10 ir 20.
- Navikas gali būti jaučiamas atliekant tiesiosios žarnos tyrimą arba matomas atliekant vaizdo tyrimą, pavyzdžiui, transrektalinį ultragarsą. Jis yra vienoje prostatos pusėje arba daugiau nei pusėje vienos prostatos pusės, arba jis yra abiejose prostatos pusėse.
- Prostata buvo pašalinta operacijos būdu, o auglys vis dar buvo tik prostatoje.

III stadijos apibūdinimas (žr. 6 pav.) [13]:

- Vėžys nėra išplitęs į netoliese esančius limfmazgius ar kitur organizme.
- PSA lygis yra virš 20 arba kitais atvejais jo vertė gali būti bet kokia.
- Vėžys gali vis dar būti tik prostatoje arba jis jau išplitęs už prostatos ribų, pavyzdžiui, į šlaplės sfinkterį (raumenį, kuris padeda kontroliuoti šlapinimąsi), tiesiąją žarną, šlapimo pūslę arba dubens sieną.

IV stadijos apibūdinimas (žr. 6 pav.) [13]:

- Vėžys gali augti arba gali neaugti į audinius, esančius šalia prostatos.
- Vėžys gali išplisti arba neišplisti į netoliese esančius limfmazgius.
- Vėžys gali išplisti į kitas kūno dalis, pavyzdžiui, tolimus limfmazgius, kaulus ar kitus organus.
- PSA lygis gali būti bet kokios vertės.



6 pav. Priešinės liaukos vėžio stadijos [14]

Nors prostatos vėžio stadija gali padėti suprasti paciento ligos sunkumą ir pavojingumą, tačiau gydytojai dabar ieško kitų būdų pasakyti, kokia yra prostatos vėžio augimo ir plitimo tikimybė, o tai taip pat gali padėti nustatyti geriausias vyro gydymo galimybes [13].

Taigi priešinės liaukos vėžys vis dar nėra aiškiai suprantama ar visais atvejais išgydoma liga ir išlieka viena iš aktualiausių problemų šių dienų medicijoje.

1.3. Internetinių svetainių kūrimas, WIX

Internetas yra pasaulinis tinklas, kuris jungia kompiuterines sistemas [15]. Tai apima kelias didelio pralaidumo duomenų linijas, kurios yra prijungtos prie pagrindinių interneto centrų, kurie platina duomenis į kitas vietas, pavyzdžiui, žiniatinklio serverius ar interneto paslaugų teikėjams [15]. Žmogus, norintis prisijungti prie interneto, turi kreiptis į interneto paslaugų teikėją. Interneto paslaugų tiekėjas yra tarpininkas tarp žmogaus ir interneto [15].

Tinklapis yra paprastas dokumentas, kurį gali rodyti naršyklė. Tokie dokumentai yra parašyti HTML kalba [16]. Tinklapis gali įterpti įvairių tipų išteklius [16]:

- stiliaus informacija (puslapio išvaizdos valdymas);
- scenarijai (suteikia puslapiui interaktyvumo);
- medija (vaizdai, garsai ir vaizdo įrašai).

Visus žiniatinklio puslapius galima pasiekti unikaliu adresu. Norint rasti puslapį reikia įvesti jo adresą į naršyklės adreso juostą [16].

Šiais technologijų laikais yra sukuriama vis daugiau įrankių, padedančių kurti internetines svetaines. Daugybė svetainių kūrėjų siūlo galimybę klientui sukurti svetainę nemokamai. Internetinės svetainės talpinamos debesyje. Viskas, ko kūrėjui reikia, yra kompiuteris ir interneto naršyklė, ji gali būti [17]:

- Google Chrome;

- Mozilla Firefox;
- Microsoft Edge ir kiti.

Robert Brandil „Website tool Tester“ puslapyje publikuojamame straipsnyje pateikiamos aukščiausiai įvertintos internetinių svetainių kūrimo platformos [17]:

1. Webnode;
2. Wix;
3. Weebly;
4. Ucraft;
5. Webstarts;
6. Jimdo;
7. Site123;
8. GoDaddy;
9. Webflow;
10. Mozello;
11. Strikingly.

Viena iš didžiausių platformų yra „Wix“. Ji turi daugiau vartotojų nei bet kuris kitas svetainių kūrėjas [17].

2006 metais buvo sukurta internetinių svetainių kūrimo programa „Wix“ [18]. „Wix“ yra debesimis paremtas įrankis, turintis milijonus vartotojų visame pasaulyje. 2009 metais jų sukurtas įrankis jau pasiekė vieną milijoną vartotojų [18]. 2014 metais vartotojų skaičius išaugo iki penkiasdešimties milijonų, o 2021 metais du šimtai milijonų vartotojų pasaulyje pradėjo naudotis jų sukurta programa [18]. Jų plati vartotojų bazė, atviras programinės įrangos kūrimo rinkinys ir neprilygstamos dizaino galimybės pritraukia vartotojų mases. „Wix“ yra klasikinis svetainių kūrėjas ir dirbtinio intelekto platforma, sujungta į vieną.

1.4. Statistinės analizės sistema, R

R yra programavimo kalba, taip pat ir statistinio skaičiavimo bei grafinio vaizdavimo aplinka [19]. Tai GNU, oficialios R programinės įrangos aplinkos, projektas, kuris yra panašus į S kalbą ir aplinką, o jį sukūrė „Bell Laboratories“ darbuotojas John Chambers ir jo kolegos [19]. R gali būti vertinama kaip kitokio tipo S kalba. Yra keletas svarbių šių programavimo kalbų skirtumų, tačiau daugelis S kalba parašytų kodų veikia ir R programos aplinkoje [19].

R teikia daugybę statistinių ir grafinių metodų. S kalba dažnai yra pasirenkama realizuoti statistinės analizės metodus, taip pat R suteikia atvirojo kodo programavimo galimybę analizuoti statistinius metodus. Vienas iš R privalumų yra tai, kad ji leidžia lengvai sukurti aukštos kokybės grafikus tinkamose vietose panaudojant matematinius simbolius ir formules [19].

R yra nemokama programinė įranga, kuri paremta laisvosios programinės įrangos fondo GNU bendrosios viešosios licencijos šaltinio kodo sąlygomis [19]. Jis kompiliuojamas ir veikia įvairiose UNIX platformose bei panašiose sistemose [19]:

- FreeBSD;
- Linux;
- Windows;
- MacOS.

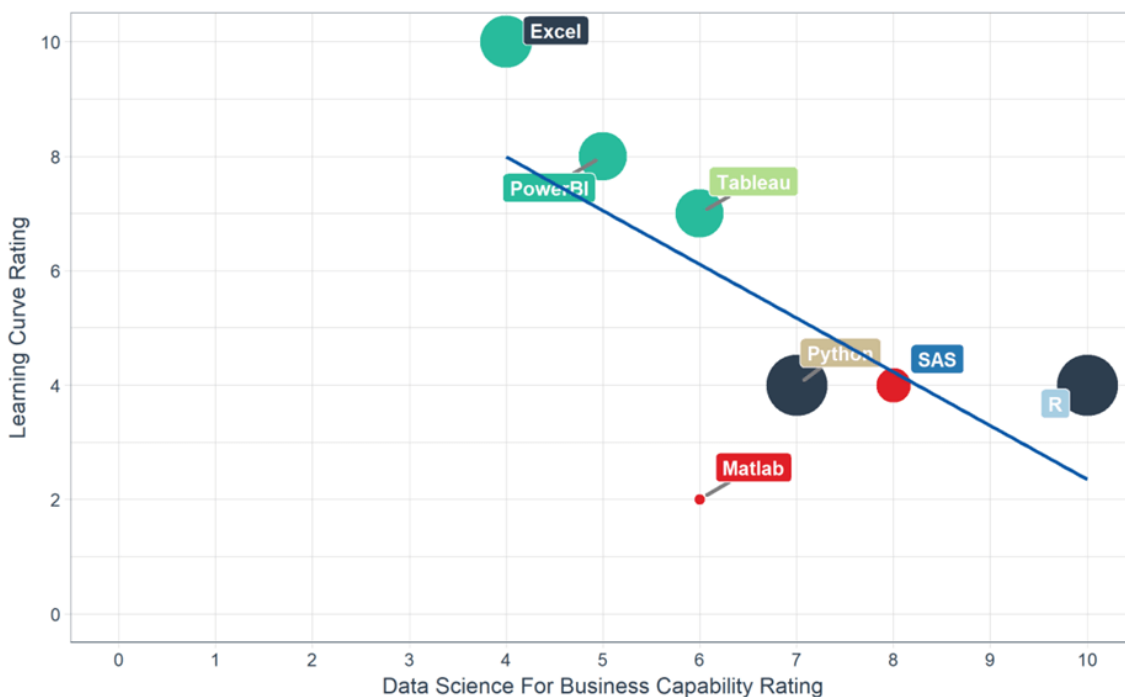
R programinė įranga gali būti palyginta su populiariais komerciniais statistikos paketais [19]:

- SAS;
- SPSS;
- Stata ir kiti.

Tačiau dar vienas R privalumas yra tai, jog ji vartotojams pasiekama nemokamai pagal nemokamos programinės įrangos licenciją [19].

2009 metų sausio mėnesį „New York Times“ paskelbė straipsnį, kuriame aprašomas R programavimo kalbos naudojimo augimas. Ją labiausiai pamėgo mokslininkai bei duomenų analitikai. Taip pat aprašoma ir jos populiarumo didėjimo priežastys ir grėsmės, kurias ji kelia komerciniams statistikos paketams [19]. 2017 metų birželio mėnesį duomenų analitikas Robert Muenchen išleido išsamesnį R ir kitų programinės įrangos paketų palyginimo straipsnį „Duomenų mokslo programinės įrangos populiarumas“ [19].

Šių dienų programinių paketų rinkoje yra daugybė lengvai prieinamų įrankių skirtų atlikti duomenų analizę. Tačiau norint išmokti naują programavimo kalbą, reikia investuoti ne tik begalę laiko, bet ir daug pastangų. Žemiau pateiktame paveiksle pavaizduota mokymosi kreivė, palyginta su verslo galimybėmis, kurias teikia įvairios programavimo kalbos (žr. 7 pav.) [20]. Norint iš turimų duomenų pateikti geriausias išvalgas, reikia laiko išmokti valdyti ir naudoti tinkamą įrankį, tokį kaip R.



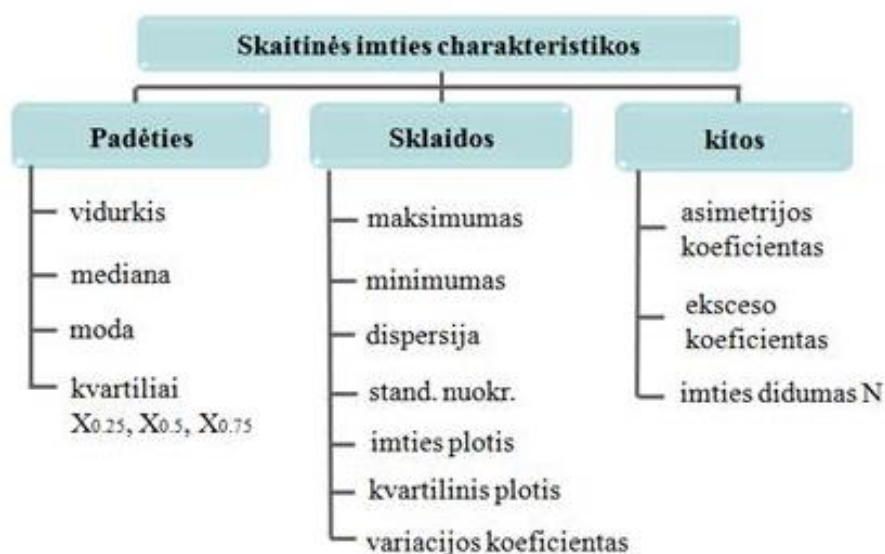
7 pav. Programinės įrangos galimybių ir mokymosi kreivė [20]

Trumpai tariant, R yra puiki priemonė studijuoti ar tirti duomenis. Išsami analizė, pavyzdžiui, grupavimas, koreliacija ar duomenų imties mažinimas bei daugelis kitų veiksmų yra atliekami naudojantis šiuo įrankiu. Tai yra viena iš svarbiausių bet kokio projekto dalių. R turi platų funkcionalumą duomenų analitikos moksle ir jos nemokama paslauga užtikrina didelę vartotojų rinką.

2. Projekto tyrimo metodai

2.1. Aprašomoji statistika

Statistika – tai mokslas, pagrįstas duomenų rinkimu, tų duomenų aprašymu, susisteminiu, analizavimu bei interpretavimu. Aprašomoji statistika yra duomenų sisteminimo ir grafinio vaizdavimo metodai [21]. Vienas iš daugelio aprašomosios statistikos privalumų yra tas, jog ji leidžia informaciją, esančią dideliuose duomenų masyvuose, užrašyti koncentruotai [21]. Čia stebimos reikšmės gali būti pateikiamos arba grafikais, arba lentelėmis, arba dažnių skirstiniais. Skiriamos kelios skaitinės imties charakteristikos – tai sklaidos bei duomenų padėties charakteristikos, taip pat charakteristikos įvertinti imties simetriškumą, asimetrijos ir eksceso koeficientai ir kita (žr. 8 pav.) [21].



8 pav. Skaitinės charakteristikos [21]

Visuotinėje lietuvių enciklopedijoje vidurkis yra apibrėžiamas kaip pagrindinė atsitiktinio dydžio skaitinė (centro) charakteristika [22]. Šis dydis apibūdina atsitiktinių dydžių reikšmių išsidėstymą, taip pat žinomą kaip svorio centras. Tai yra reikšmė, apie kurią atsitiktinis dydis mažiausiai išsisklaidęs vidutinio kvadratinio nuokrypio prasme [22]. Vidurkis yra apskaičiuojamas formule (žr. 1 formulę):

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n} \quad (1)$$

čia $x_1 + x_2 + \dots + x_n$ yra tiriamo kintamojo imties skaitinės reikšmės, n yra reikšmių skaičius imtyje.

Moda yra atsitiktinio dydžio pasiskirstymo charakteristika, kitaip tariant, tai yra atsitiktinio dydžio pasiskirstymo tankio $p(x)$ maksimumo taškas [22]. Trumpai tariant, tai dažniausiai pasitaikanti reikšmė imtyje. Pagal modų skaičių yra skiriami šių tipų skirstiniai [22]:

1. unimodinis (viena moda);
2. bimodinis (dvi modos);
3. multimodinis (daug modų).

Mediana (lot. mediana – vidurinė) yra atsitiktinio dydžio pasiskirstymo skaitinė charakteristika, arba kitaip tai yra vidutinė (centrinė) atsitiktinio dydžio reikšmė [22]. Mediana dalija populiaciją, variacinę eilutę arba tikimybinį pasiskirstymą į dvi lygias dalis [22].

Kvantiliai yra duomenų padėtį apibūdinančios charakteristikos, o kvartilai yra kvantiliai, kurie dalija variacinę duomenų eilutę į keturias lygias dalis [23].

„Visos charakteristikos, išskyrus modą, gali būti skaičiuojamos tik kiekybiniais duomenimis“ [21].

Statistinis dažnis yra rodiklis, parodantis, kaip dažnai tam tikra reikšmė pasitaiko duomenų eilutėje, o dažnių lentelė parodo, kaip dažnai kiekviena iš nurodytų reikšmių pasitaiko ar patenka į tam tikrą reikšmių intervalą [23].

Imties dydis yra imtyje egzistuojančių elementų (pvz., tiriamųjų, pacientų ar pan.) skaičius, kitaip dar vadinamas imties tūriu [23].

Maksimumas yra didžiausia reikšmė duomenų imtyje, o minimumas yra mažiausia reikšmė duomenų imtyje.

Dispersija yra duomenų sklaida aplink vidurkį. Ji apskaičiuojama formule (žr. 2 formulę) [24]:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2)$$

čia x_i yra tiriamo kintamojo imties skaitinės reikšmės, n yra reikšmių skaičius imtyje, o \bar{x} yra imties vidurkis.

Standartinis nuokrypis yra apskaičiuojamas formule (žr. 3 formulę) [24]:

$$s = SD = \sqrt{s^2} \quad (3)$$

čia s^2 yra dispersija.

Asimetrijos koeficientas yra naudojamas įvertinti duomenų asimetriją ir apskaičiuojamas formule (žr. 4 formulę) [24]:

$$g_1 = \frac{n \sum_{i=1}^n (x_i - \bar{x})^3}{(n-1)(n-2)s^3} \quad (4)$$

čia s yra standartinis nuokrypis, x_i yra tiriamo kintamojo imties skaitinės reikšmės, n yra reikšmių skaičius imtyje, o \bar{x} yra imties vidurkis. Kai $g_1 > 0$, tada asimetrija yra teigiama, o kai $g_1 < 0$ tada asimetrija yra neigiama, o kai $g_1 = 0$, simetrija yra ideali.

Atliekant aprašomąją statistiką reikia nustatyti kintamųjų tipus. Nustačius tipus kategoriniams nominaliesiems duomenims, yra pateikiama [23]:

- imties dydis;
- dažnių lentelė.

O tolydiems duomenims pateikiama [23]:

- imties dydis;
- duomenų centras (tai vidurkis ar mediana);
- duomenų sklaida (standartinis nuokrypis, kvantiliai ir kita).

Tuo atveju, kai nustatomi ranginiai arba skaitiniai diskretieji duomenys, jiems gali tikti ir tolydžiųjų, ir diskrečiųjų duomenų aprašymo metodai [23]. Tokiu atveju tenka įvertinti, į ką konkretūs kintamieji yra panašesni – į tolydžiuosius (t.y., kiekybinius) ar į diskrečiuosius (t.y., nominaliuosius). Atsižvelgiant į tai, yra parenkami aprašomosios statistikos metodai.

Ranginių duomenų metodai[23]:

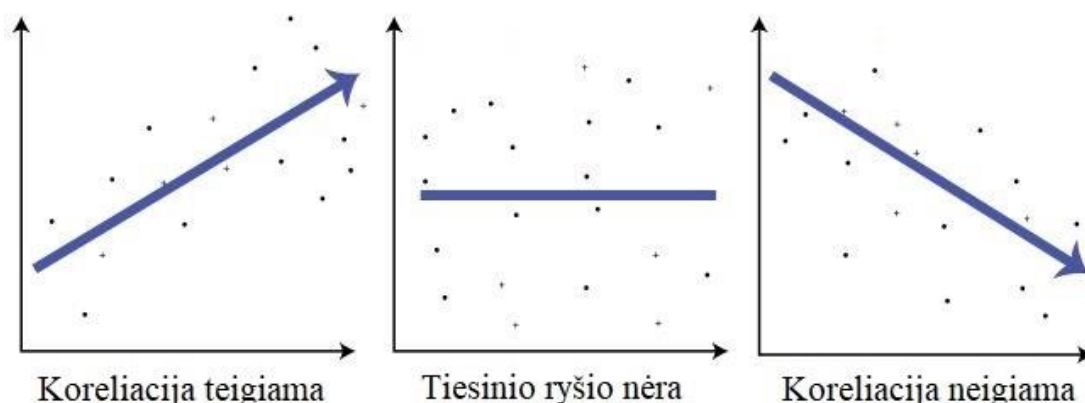
- imties dydis;
- dažnių lentelė;
- mediana;
- kiti centrą ir sklaidą aprašantys dydžiai naudojami tik tuo atveju, jeigu jie turi prasmę.

Diskrečiųjų skaitinių duomenų metodai [23]:

- imties dydis;
- duomenų centras;
- duomenų sklaida;
- dažnių lentelė.

2.2. Koreliacinė analizė

Medicinoje dažnai reikia daryti išvadas apie dviejų ar daugiau kintamųjų ryšį. Tokiems ryšiams nustatyti yra pasitelkiama koreliacinė analizė. Koreliacija, dar vadinama koreliacine analize, yra terminas, kuris vartojamas žymėti dviejų (ar daugiau) kiekybinių kintamųjų ryšius [25]. Iš esmės ši analizė yra pagrįsta tiesinio ryšio tarp kiekybinių kintamųjų prielaida [25]. Panašiai kaip dvejetainių kintamųjų asociacijos matai, šis dydis matuoja asociacijos tarp kintamųjų stiprumą arba mastą ir jo kryptį [25]. Koreliacijos koeficientas yra galutinis koreliacijos analizės rezultatas, o jo vertės gali būti nuo -1 iki $+1$. Teigiamas koreliacijos koeficientas $+1$ rodo, jog du kintamieji yra tarpusavyje susiję teigiamai, tiesiškai. Kai koreliacijos koeficientas yra lygus -1 , tai rodo, kad du kintamieji yra susiję neigiamai, tiesiškai. Tik tuo atveju, kai koreliacijos koeficientas yra lygus nuliui, tai rodo, kad tarp dviejų tiriamų kintamųjų nėra tiesinio ryšio (žr. 9 pav.).



9 pav. Koreliacijos pavyzdys [25]

Paprastai statistikoje yra matuojamos šios koreliacijų rūšys [26]:

- Pirsono koreliacija;
- Kendalo rango koreliacija;
- Spirmano koreliacija.

Pirsono koreliacija yra plačiausiai naudojama koreliacijos statistika matuoti santykį tarp tiesiškai susijusių kintamųjų laipsnių ir ji yra apskaičiuojama naudojantis formule (žr. 5 formulę) [26]:

$$r_{xy} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}} \quad (5)$$

čia r_{xy} – Pirsono koreliacijos koeficientas tarp x ir y ; n – stebėjimų skaičius; x_i – x vertė (i-tajam stebėjimui); y_i – y vertė (i-tajam stebėjimui).

Pirsono koreliacijoje abu kintamieji turėtų būti normaliai pasiskirstę (paprastai taip pasiskirstę kintamieji turi varpo formos kreivę). Kitos prielaidos apima tiesiškumą ir homoskedastiškumą. O linijiškumas leidžia daryti prielaidą, kad tarp abiejų kintamųjų yra tiesi linija, o homoscedastika – tai, kad duomenys yra vienodai paskirstyti apie regresijos tiesę [26].

Kendalo (angl. Kendall's tau) rango koreliacija yra neparametrinis testas ir jis matuoja priklausomybės tarp dviejų kintamųjų stiprumą. Ši koreliacija yra lengviau interpretuojama negu Spirmeno ir generalinėje aibėje tiksliau įvertina tikrąją koreliaciją, todėl geriau generalizuoja ir labiau tinka nedidelėms imtims, tokioms kuriose yra didelis kiekis pasikartojančių reikšmių [27]. Tarkime, dviejų kintamųjų a ir b imtyje, kur kiekvienas imties dydis yra n , žinome, kad bendras porų su b skaičius yra $n(n-1)/2$. Kendall rango koreliacijos vertei apskaičiuoti naudojama formulė (žr. 6 formulę) [26]:

$$\tau = \frac{n_c - n_d}{\frac{1}{2}n(n-1)} \quad (6)$$

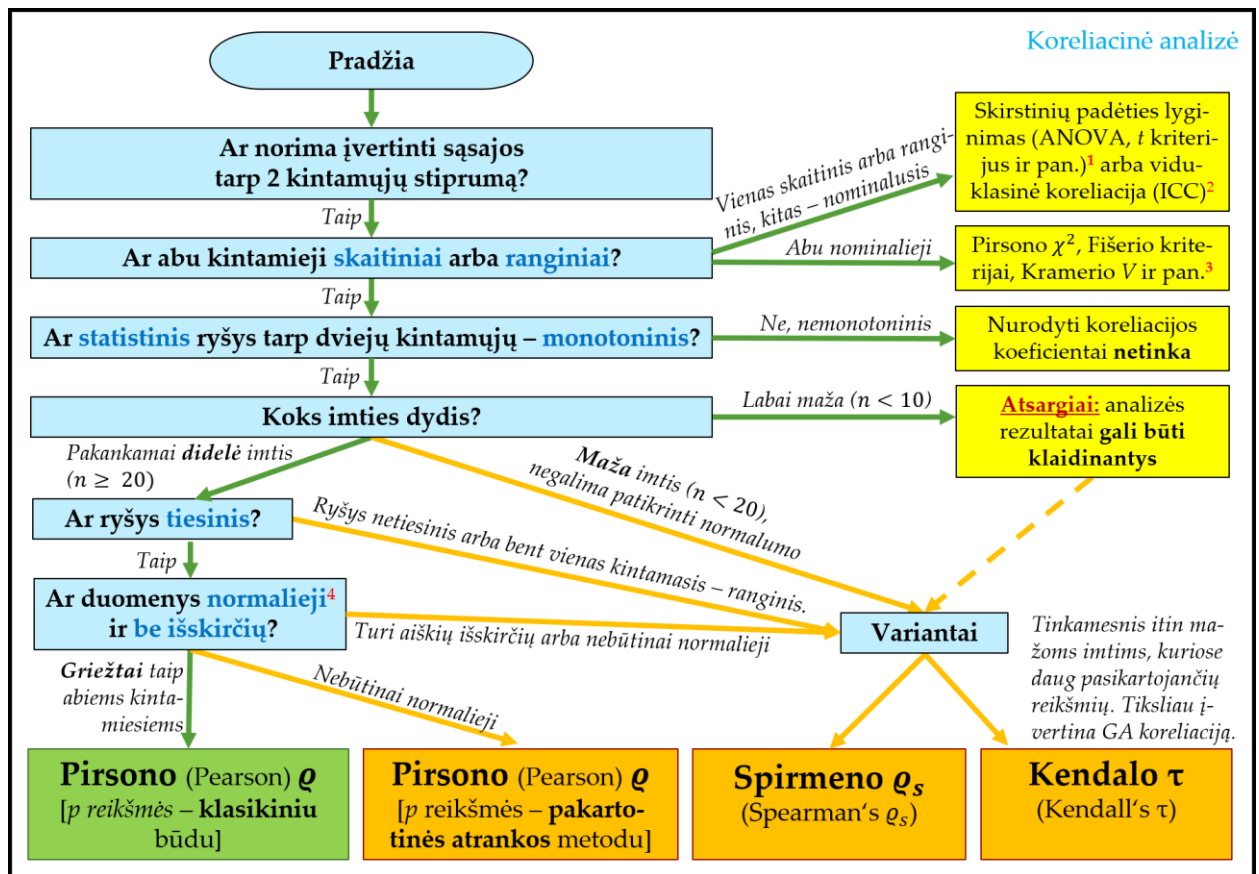
čia n_c – suderintoji imtis; n_d – skirtingumų imtis.

Spirmeno rango koreliacija yra neparametrinis testas ir jis yra naudojamas matuojant dviejų kintamųjų sąsajos laipsnį [26]. Spirmeno rango koreliacija neturi jokių prielaidų apie duomenų pasiskirstymą ir yra tinkama koreliacinei analizei atlikti, kai kintamieji matuojami skalėje, kuri yra ordinali. Ji apskaičiuojama naudojantis formule (žr. 7 formulę) [26]:

$$\rho = \frac{\delta \sum d_i^2}{n(n^2-1)} \quad (7)$$

čia ρ – Spirmeno koreliacijos koeficientas; d_i – skirtumas tarp atitinkamų kintamųjų rangų; n – stebėjimų skaičius.

Tinkamai pasirinkti koreliacijos metodą yra labai svarbu. Visų pirma, reikia įvertinti kintamųjų tipus, tada imties dydį, nustatyti, ar yra tiesinis ryšys tarp kintamųjų ir ar duomenys normalieji ir ar juose yra išskirčių (žr. 10 pav.) [27].

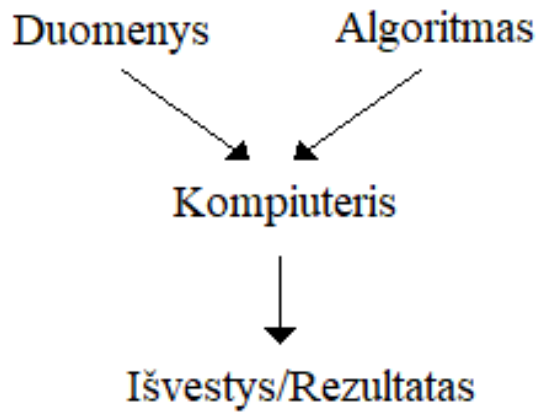


10 pav. Schema, rekomenduojanti, kurį koreliacijos koeficientą pasirinkti [27]

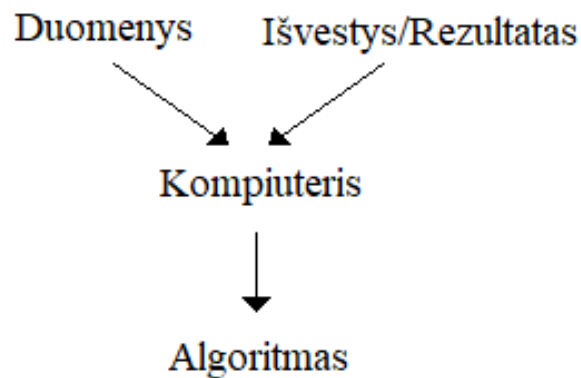
2.3. Mašininis mokymasis ir dirbtiniai neuroniniai tinklai

XXI amžiuje dirbtinis intelektas tapo ypač populiaria tema tiek mokslo bendruomenėje, tiek už jos ribų. Gausybė straipsnių bei žurnalų aptarė mašininio mokymosi, gilaus mokymosi ir dirbtinio intelekto temas. Visi šie terminai yra glaudžiai susiję.

Mašininis mokymasis yra sritis, orientuota į dirbtinio intelekto mokymąsi, kuriant algoritmus, kurie geriausiai atspindi duomenų rinkinį [28]. Skirtingai nuo klasikinio programavimo, kai algoritmą galima aiškiai užkoduoti naudojant žinomas ypatybes, mašininis mokymasis naudoja duomenų pogrupius generuoti algoritmą [28]. Šiuose pogrupiuose gali būti naudojamos funkcijų ir svorio kombinacijos. (žr. 11, 12 pav.).



11 pav. Klasikinio programavimo realizacija



12 pav. Mašininio mokymosi realizacija

Programuojant klasikiniu būdu kompiuteriui yra pateikiamas duomenų rinkinys ir algoritmas. Po to algoritmas informuoja kompiuterį, koku būdu reikia valdyti duomenų rinkinį kuriant išvestis (žr. 11 pav.). O mašininis mokymasis skiriasi tuo, kad kompiuteriui yra pateikiamas duomenų rinkinys ir išvestys arba kiti gauti rezultatai. Turėdamas ir įvestis, ir rezultatus kompiuteris „apsimokina“ ir sugeneruoja algoritmą, kuris apibūdina jų tarpusavio ryšį ir išvestis (žr. 12 pav.). Šio tipo algoritmas gali būti naudojamas išvadoms apie būsimus duomenų rinkinius.

Mašininis mokymasis turi keturis dažniausiai naudojamus mokymosi metodus, kurių kiekvienas yra naudingas sprendžiant įvairias užduotis [28]:

- prižiūrimas;
- neprižiūrimas;
- pusiau prižiūrimas;
- pastiprintas mokymasis.

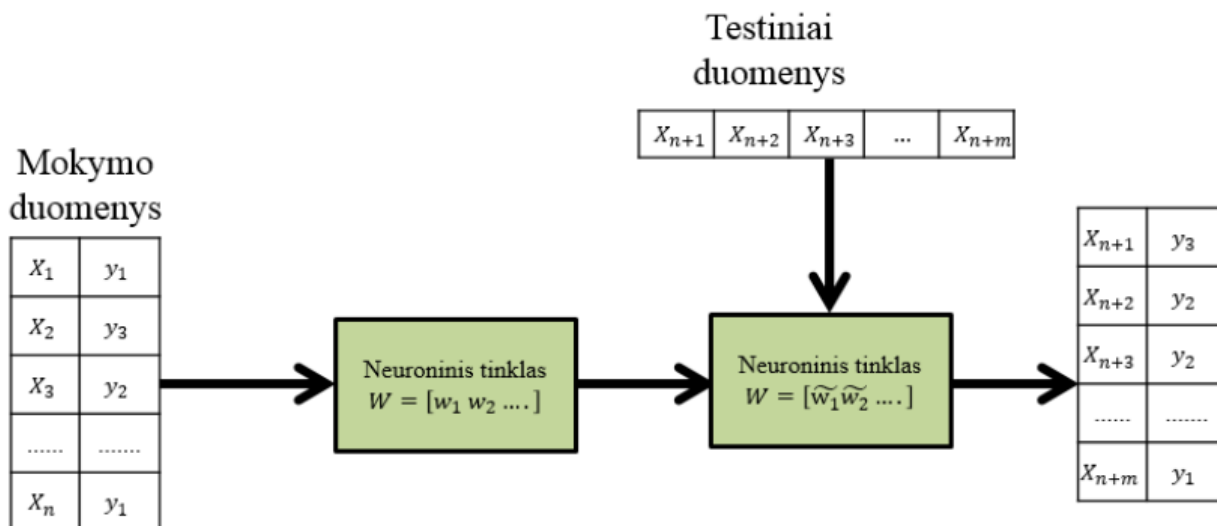
Medicinoje naudojama daug mašininio mokymosi algoritmų [28]:

- tiesinė regresija;
- logistinė regresija;
- sprendimų medžiai ir atsitiktiniai miškai;
- klasikinė oftalmologija;
- dirbtiniai neuroniniai tinklai.

Dirbtinis neuroninis tinklas yra mašininio mokymosi algoritmas. Jis paremtas biologinių neuroninių tinklų logika [28]. Kiekviename neuroniniame tinkle yra mazgų, kurie nustatytais ryšiais bendrauja su kitais mazgais. Dirbtiniu neuronu yra vadinama funkcija, turinti baigtinį skaičių įvesties reikšmių bei vienintelę išvesties reikšmę [29]. Dirbtiniai neuronai kurie yra susijungę tarpusavy yra vadinami dirbtiniu neuroniniu tinklu. Sluoksnių ir juose esančių neuronų skaičius priklauso nuo sprendžiamo uždavinio tipo ir jo sudėtingumo [29]. Neuroniniuose tinkluose objektai yra aprašomi vektoriumi (žr. 8 formulę.) [29]:

$$X = [x_1 \ x_2 \ x_3 \ \dots \ x_n] \quad (8)$$

Aibė $Y = [y_1 \ y_2 \ \dots \ y_m]$ aprašo visas y_i klases, kurioms gali priklausyti objektas X . Mokymosi fazėje imami apmokymo duomenys. Optimizuojant algoritmus parenkami neuroninio tinklo koeficientai w_1, w_2, \dots, w_k taip, kad klasifikavimo funkcija K pavaizduotų mokymosi etapo objektus X į klasę y , kuriai jie priklauso [29]. Testavimo etape yra imama kita porų aibė (X_i, y_j) . Taip apskaičiuojama, kiek kartų objektai yra teisingai suklasifikuoti [29]. Taip yra įvertinamas klasifikatoriaus tikslumas. Jeigu tikslumas tenkina, tai neuroninis tinklas paruoštas darbui, kitu atveju pradiniai duomenys yra koreguojami. Keičiamas pradinių reikšmių parinkimas ir iš naujo mokomas neuroninis tinklas [29] (žr. 13 pav.).

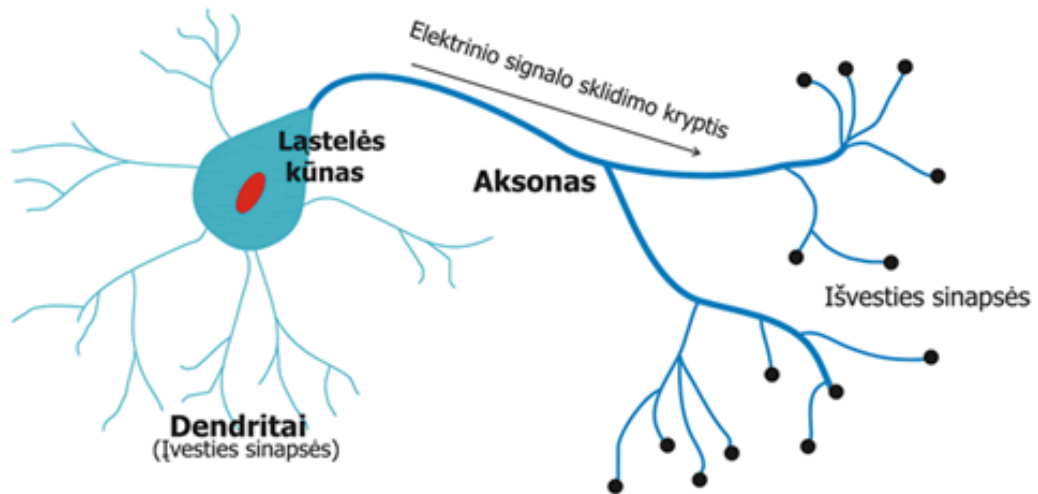


13 pav. Neuroninio tinklo mokymo ir testavimo fazės [29]

Statistinė programinės įrangos kalba R su neuralnet biblioteka naudojama konstruojant skirtingus dirbtinius neuroninius tinklus [30]. Šio paketo naudojimas leidžia sukurti dirbtinius neuroninius tinklus, turinčius skirtingą paslėptų sluoksnių ir neuronų skaičių. Tačiau šiam metodui yra labai svarbu [30]:

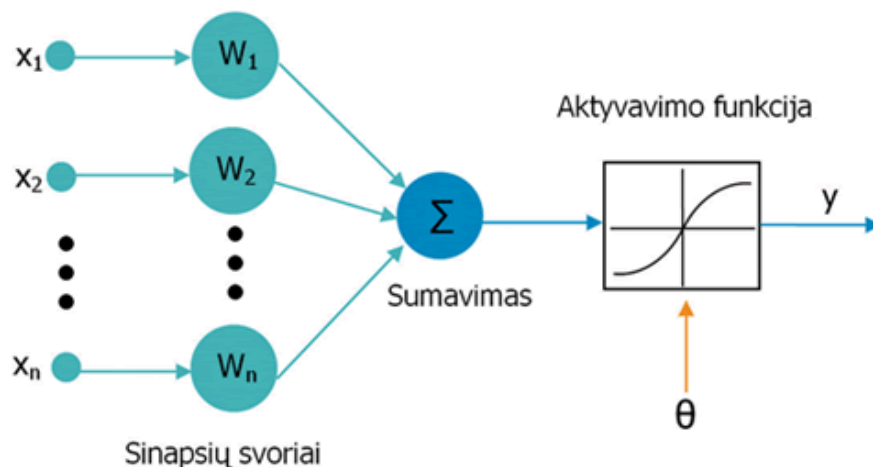
- teisingai parinkti įvesties parametrus;
- duomenis apdoroti iš anksto;
- normalizuoti duomenis.

Biologinį neuronų tinklą sudaro nervinės ląstelės, kurios vadinamos neuronais [30]. Kiekvienas neuronas turi įėjimus, vadinamus sinapsėmis, ir išėjimus, vadinamus aksonais, su kuriais neuronai yra susieti (žr. 14 pav.) [30].



14 pav. Biologinio neurono sandara [31]

Informacija, tai yra elektriniai impulsai, kuriuos neuronas gauna iš įvesčių, yra apdorojami [30]. Taip yra nustatoma neurono būklė, kuri gali būti sužadinta arba nesužadinta [30]. Per išvestis ši būseną perduodama kitiems neuronams. Kai kompiuteryje yra imituojamas biologinis neuroninis tinklas, gaunamas vadinamasis dirbtinis neuroninis tinklas. Taip pat kaip ir biologinį neuronų tinklą, šį tinklą sudaro neuronai, turintys ryšį vienas su kitu [30]. Tokie neuronai yra išdėstyti sluoksniais. Jei kiekvienas sluoksnio neuronas turi ryšius su visais kito sluoksnio neuronais, tada toks tinklas yra vadinamas pilnai suderintu [30]. Iš gaunamos informacijos neuronas nustato savo būseną naudodamasis funkcija, paprastai susumuodamas įvestis ir kitą mastelį su aktyvacijos funkcija [30]. Ryšiai tarp neuronų gali būti įvertinti skirtingais svoriais, o kiekviename sluoksnyje yra įtrauktas poslinkio svoris (žr. 15 pav.) [30].



15 pav. Dirbtinio neurono sandara [31]

Dirbtinį neuroną sudaro trys pagrindiniai komponentai [31]:

- svoriai;
- slenksčiai;
- viena aktyvacijos funkcija.

Svorio koeficientai $W = [w_1, w_2, \dots, w_n]$ rodo atskirų įvesčių, kurios aprašomos vektoriumi $X = [x_1, x_2, x_3, \dots, x_n]$, stiprumą. Neuroninė jungtis WX yra gaunama dauginant kiekvienos įvesties signalą iš svorio koeficiento. Kai svorio koeficientas yra teigiamas, tada XW sužadina signalą išvestyje y , o jeigu svorio koeficientas yra neigiamas, XW slopina išvesties signalą. Neurono y išvesties aktyvavimą veikia vidinis neurono slenkstis (žr. 15 pav. ir 9 formulę) [31]:

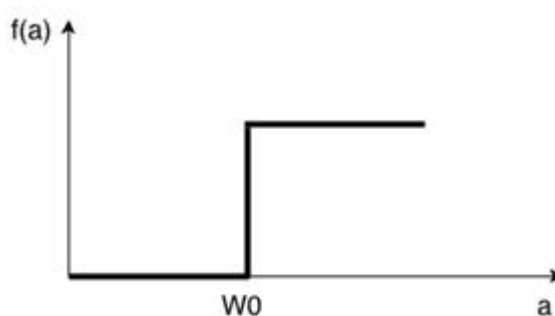
$$y = f(a) = f(\sum_{i=1}^n XW - \Theta) \quad (9)$$

čia y – aktyvacijos funkcija. W – sinapsių svorio koeficientai. X – įvesčių stiprumo rodiklis.

Aktyvacijos funkcijos reikšmė gali įgyti tokias reikšmes (žr 10 formulę) [32]:

$$f(a) = \begin{cases} 1, & \text{jei } a \geq w_0 \\ 0, & \text{jei } a < w_0 \end{cases} \quad (10)$$

čia w_0 yra slenkščio reikšmė (žr. 16 pav.).



16 pav. Slenkstinė funkcija [32]

Aktyvacijos funkcija yra matematinė operacija su išvesties signalu. Jos sudėtingumas priklauso nuo sprendžiamo uždavinio. Galimos įvairios aktyvacijos funkcijos [31]:

- hiperbolinė tangento;
- sigmoidinė;
- tiesinė;
- slenkščio ir kita.

Vienas iš mokymosi su mokytoju algoritmų yra perceptronas, priklausantis tiesioginio sklidimo neuroniniams tinklams [32]. Perceptrono mokymas svorius keičia taip, kad „pateikus į tinklą vektorių X_i , skirtumas tarp išėjimo vektoriaus Y_i ir norimo vektoriaus T_i būtų kiek galima mažesnis“ (žr. 11 formulę) [32]:

$$E(W) = \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^d (Y_i - T_i)^2 \rightarrow \min \quad (11)$$

čia $Y_i = (y_{i1}, \dots, y_{id})$. $T_i = (t_{i1}, \dots, t_{id})$. d yra išėjimų skaičius. m yra įėjimo vektorių skaičius. $W = \{w_{jk}, j = 1, \dots, d, k = 0, \dots, n\}$. w_{jk} yra jungties, iš k -tojo įėjimo į j -tąjį neuroną, svoris.

R programavimo įrankyje esantis paketas neuralnet yra sukurtas mokyti daugiasluoksnius perceptronus [30]. Naudojantis šiuo paketu galima susikurti daugiasluoksnius perceptronus, kurie tinka modeliuoti funkcinis santykius [30].

Neuroniniame tinkle sinapsė gali būti sujungta tik su tolesniais lygiais. Įvesties sluoksnis susideda iš visų lydinčių kintamųjų atskiruose neuronuose, o išvesties sluoksnis susideda iš atsako kintamųjų

[30]. Tarpusavio sluoksniai yra vadinami paslėptaisiais. Įvesties sluoksnyje ir paslėptuose sluoksniuose yra pastovus neuronas, jis yra poslinkis. Neuralnet yra oreantuotas į prižiūrimus mokymosi algoritmus. Jiems būdingas duotos išvesties naudojimas (mokymo stadija) kuris yra lyginamas su apskaičiuota išvestimi. Neuroninio tinklo parametrai yra jo svoriai, kurie paprastai inicijuojami atsitiktinėmis reikšmėmis [30]. Neuroninio tinklo mokymo metu yra atliekami šie veiksmai [30]:

- apskaičiuojamos išvestys duotoms įvestims ir dabartiniams svoriams;
- apskaičiuojama klaidos funkcija;
- apskaičiuojamas skirtumas tarp prognozuojamos ir stebimos išvesties;
- svoriai yra pritaikomi pagal mokymosi algoritmo taisyklę.

Šis procesas yra sustabdomas tuo atveju, jeigu patenkinamas iš anksto nurodytas kriterijus [30]. Po to taikomas atgalinio sklidimo algoritmas. Atgalinio sklidimo algoritmas yra pagrįstas neuroninio tinklo svorio modifikavimu todėl, kad būtų nustatytas lokalus klaidos funkcijos minimumas [30]. Dėl šios priežasties, atsižvelgiant į svorius, yra apskaičiuojamas klaidos funkcijos gradientas [30].

2.4. Daugialypė tiesinė regresija

Matematikoje regresija leidžia, naudojant vieno ar daugiau kintamųjų reikšmes X_1, X_2, \dots, X_i , prognozuoti kito kintamojo Y vertę. Tai statistinis modelis [33]. Daugialypės tiesinės regresijos modelio lygtis (žr. 12 formulę) [33]:

$$Y = a + (\sum b_i X_i) + \varepsilon \quad (12)$$

čia Y – priklausomas kintamasis, X_i – nepriklausomi kintamieji (regresoriai), a – koeficientas (laisvasis narys), b_i – krypties koeficientai (svertiniai), ε – liekamoji paklaida (angl. residual)

Sąlygos duomenims modelio naudojimui [34]:

- Paklaidos ε pasiskirstę normaliai.
- Vidurkiai $E\varepsilon_i = 0$.
- (Homoskedastiškumas) dispersijos $D\varepsilon_i = \sigma^2$.
- ε_i yra nepriklausomi.
- Išskirčių duomenyse nėra
- Multikolinearumo tarp X_i nėra.

3. Projekto tyrimo rezultatai ir aptarimas

3.1. Duomenų rinkinio parengimas

Šiame projekte naudoti duomenys buvo surinkti ir pateikti 2020 metais Lietuvos Sveikatos mokslų universiteto, Urologijos klinikos skyriaus vadovo, profesoriaus Daimanto Milono. Pradinį duomenų rinkinį sudarė 2423 pacientų tyrimų duomenys ir 12 skirtingų atributų (žr. 2 lentelę).

2 lentelė. Pradinių duomenų atributai ir jų reikšmės

Atributas	Reikšmė
Nr	Nuasmeninto paciento duomenų eilutės numeris
Amzius	Paciento amžius metais
PSAI	Prostatos specifinis antigenas
GLpoRP_ISUP	Prostatos vėžio diferenciacija operacinėje medžiagoje. Tai išvešėjusių limfmazgių skaičius 1, 2, 3, 4, 5 arba daugiau, nustatytas operuojant.
pT	Patologinė vėžio stadija. Čia 0 – pradinė stadija, vėžys nematomas iš tomografijos. 1 – vėžys formuojasi vienoje vietoje. Galimas chirurginis šalinimas, jei vėžys pakankamai mažas. 2 – vėžys vienoje vietoje stipriai išsiplėtęs. Galima gydyti chemoterapija arba chirurgiškai.
LNI01	Limfmazgių būklė. Čia 0 – limfmazgiai švarūs. 1 – limfmazgiai pažeisti vėžiu. 2 – limfmazgiai nešalinti, būklė nežinoma.
R1	Pašalintos prostatos būklė. Čia 0 – chirurgiškai pašalintos prostatos kraštai yra švarūs. 1 – chirurgiškai pašalintos prostatos kraštai yra su vėžinėmis ląstelėmis. 2 – nėra aišku.
SURVIVAL	Stebėjimas mėnesiais iki mirties arba iki paskutinio apsilankymo.
OS	Mirtis dėl bet kurios priežasties, taip pat ir vėžio. Čia 0 – Ne (pacientas gyvas). 1 – Taip (pacientas miręs).
CSS	Mirtis nuo vėžio. Čia 0 – Ne (pacientas gyvas). 1 – Taip (pacientas miręs).
Death_other_causes	Mirtis ne dėl vėžio. Čia 0 – Ne (pacientas gyvas). 1 – Taip (pacientas miręs).
Age64vs65	Amžiaus grupės. Čia 0 – iki 64 metų amžiaus (imtinai). 1 – daugiau negu 65 metų amžiaus.

Iš visos duomenų imties 2410 respondentų duomenų buvo tinkami vykdyti tolesnę analizę. Iš 13 pacientų 10 turėjo po vieną tuščią atributą, vienas turėjo keturis tuščius atributus ir du pacientai turėjo visus 12 tuščių atributų. Toliau buvo analizuojami pjūviai tų pacientų, kurie turėjo visus užpildytus atributus. Respondentai, turėję vieną ir daugiau tuščių atributų, buvo šalinami iš duomenų imties. Taigi, atlikus pašalinimą liko 2410 pacientų, kurie dalyvavo tolesniame tyrime. Tai sudarė 99,46 % pirminės duomenų imties (žr. 3 lentelę).

3 lentelė. Duomenų failo tuščių atributų skaičius

Tuščių atributų skaičius	Pacientų kiekis	Procentalus pacientų kiekis
0	2410	99,463 %
1	10	0,413 %
4	1	0,041 %
12	2	0,083 %

3.2. Statistinės analizės modelis

Aprašomoji statistika atlikta naudojant 2410 pacientų duomenis. „Nr“, „Amzius“, „PSAI“ ir „SURVIVAL“ buvo tranktuojami kaip skaitinio tipo kintamieji, o likusieji kaip ranginio tipo kintamieji (žr. 17 pav.).

```
Classes 'tbl_df', 'tbl' and 'data.frame':      2410 obs. of  12 variables:
 $ Nr          : num  1 2 3 4 5 6 7 8 9 10 ...
 $ Amzius      : num  67 68 69 69 71 72 57 54 47 68 ...
 $ PSAI        : num  21.8 6.3 8.81 6.5 12 ...
 $ GLpoRP_ISUP : Factor w/ 5 levels "1","2","3","4",...: 4 1 1 1 3 1 1 2 1 5 ...
 $ pT          : Factor w/ 3 levels "0","1","2": 3 1 1 1 1 1 3 3 1 1 ...
 $ LNI01       : Factor w/ 3 levels "0","1","2": 1 1 1 1 1 1 1 1 1 1 ...
 $ R1          : Factor w/ 3 levels "0","1","2": 2 1 1 1 1 1 2 2 2 1 ...
 $ SURVIVAL    : num  100 189 83 189 7 138 203 203 174 72 ...
 $ OS          : Factor w/ 2 levels "0","1": 2 1 2 1 2 2 1 2 1 2 ...
 $ CSS         : Factor w/ 2 levels "0","1": 2 1 2 1 1 1 1 2 1 1 ...
 $ Death_other_causes: Factor w/ 2 levels "0","1": 1 1 1 1 2 2 1 1 1 2 ...
 $ Age64vs65   : Factor w/ 2 levels "0","1": 2 2 2 2 2 2 1 1 1 2 ...
```

17 pav. Programinės įrangos R, funkcijos „str“ rezultatai apie atributus

Apskaičiuota kiekvieno skaitinio tipo kintamojo mažiausia reikšmė (minimumas), pirmasis kvartilis, mediana, vidurkis, trečiasis kvartilis ir didžiausia reikšmė (maksimumas). Kiekvieno faktoriaus tipo atributui apskaičiuotas atsakymų pasikartojimų kiekis (žr. 4, 5 lenteles ir 3 priedą).

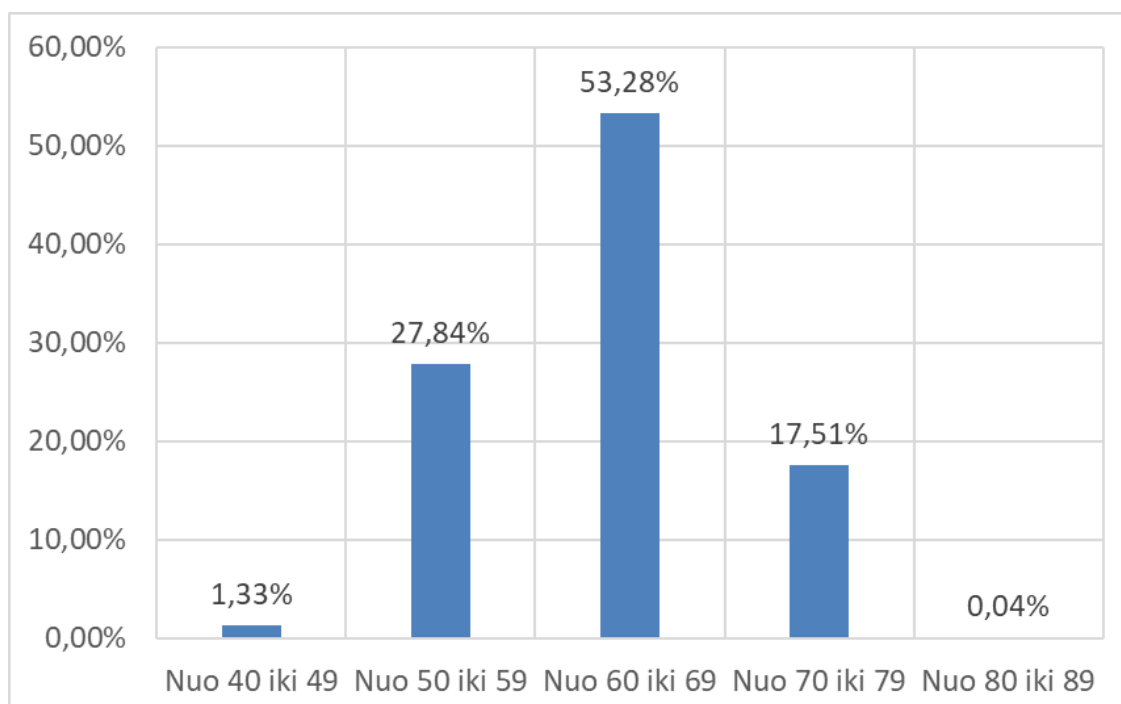
4 lentelė. Aprašomosios statistikos rezultatai, skaitinio tipo kintamieji

Atributo pavadinimas	Minimumas	Pirmasis kvartilis	Mediana	Vidurkis	Trečiasis kvartilis	Maksimumas
Nr	1,00	603,20	1205,50	1205,50	1807,80	2410
Amzius	40,00	59,00	64,00	63,19	68,00	87,00
PSAI	0,44	4,70	6,30	8,25	9,47	98,40
SURVIVAL	1,00	50,00	85,00	89,81	125,00	206,00

5 lentelė. Aprašomosios statistikos rezultatai, ranginio tipo kintamieji

Atributo pavadinimas	Galimos atributo reikšmės pasikartojimų kiekis					
	0	1	2	3	4	5
GLpoRP_ISUP	–	630	1283	250	109	138
pT	1565	661	184	–	–	–
LNI01	710	83	1617	–	–	–
R1	1606	698	106	–	–	–
OS	2060	350	–	–	–	–
OSS	2354	56	–	–	–	–
Death_other_causes	2116	294	–	–	–	–
Age64vs65	1306	1104	–	–	–	–

Toliau pacientai suskirtomi į amžiaus grupes. Jos išdėstomos dešimties metų intervalais taip, kad visa duomenų aibė patenka į tam tikrą amžiaus grupę. Patikrinamas procentalus pacientų pasiskirstymas pagal amžiaus grupes. Iš to matyti, kad daugiausiai – net 53,28 % pacientų – pateko į amžiaus grupę nuo 60 iki 69 metų. Kiek mažiau – 27,84 % pacientų – pateko į amžiaus grupę nuo 50 iki 59 metų. Į amžiaus grupę nuo 70 iki 79 metų pateko 17,51 % pacientų. Be to, labai maža dalis – tik 1,33 % pacientų – pateko į amžiaus grupę nuo 40 iki 49 metų ir vos 0,04 % į amžiaus grupę nuo 80 iki 89 metų (žr. 18 pav.). Šie rodikliai parodė, jog daugiausia serga asmenys, kurių amžius yra tarp 50 ir 79 metų.

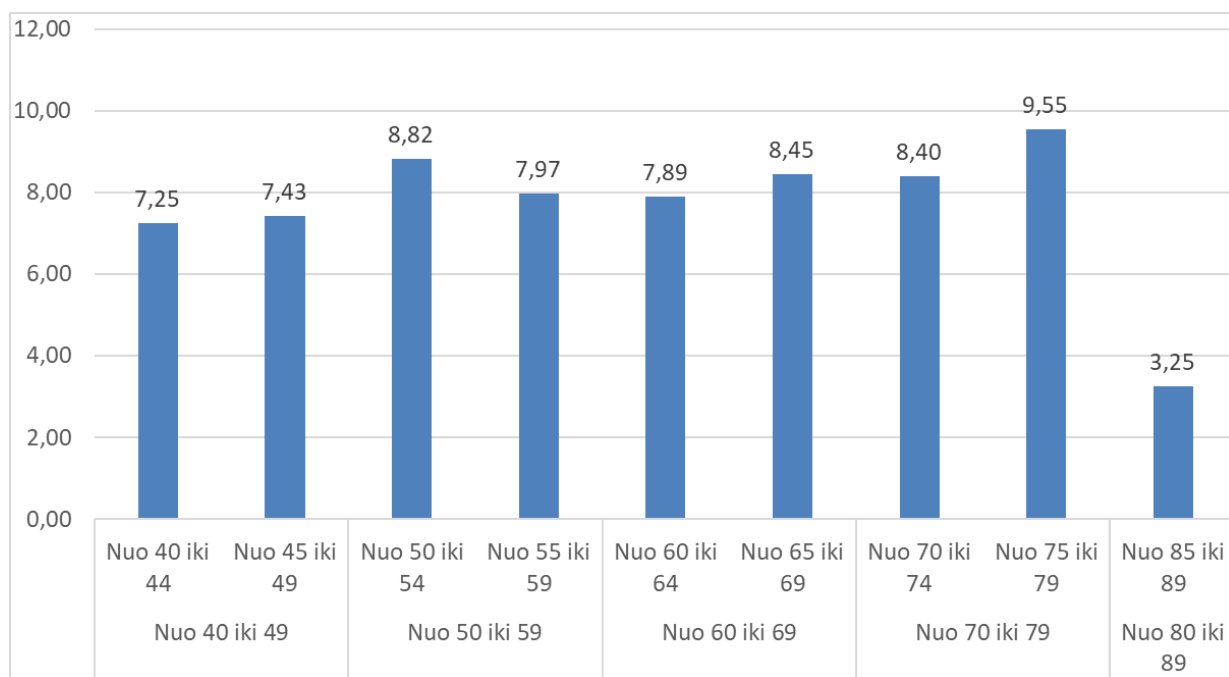


18 pav. Procentalus pacientų pasiskirstymas pagal amžiaus grupes

Tiriant skirtingo amžiaus žmonių prostatos specifinio antigeno rodiklį, duomenų aibė suskirstyta pagal amžiaus grupes, kurių intervalas yra penkeri metai. Visuose amžiaus intervaluose apskaičiuojamas prostatos specifinio antigeno rodiklio vidurkis. Iš gautų rezultatų matyti, jog

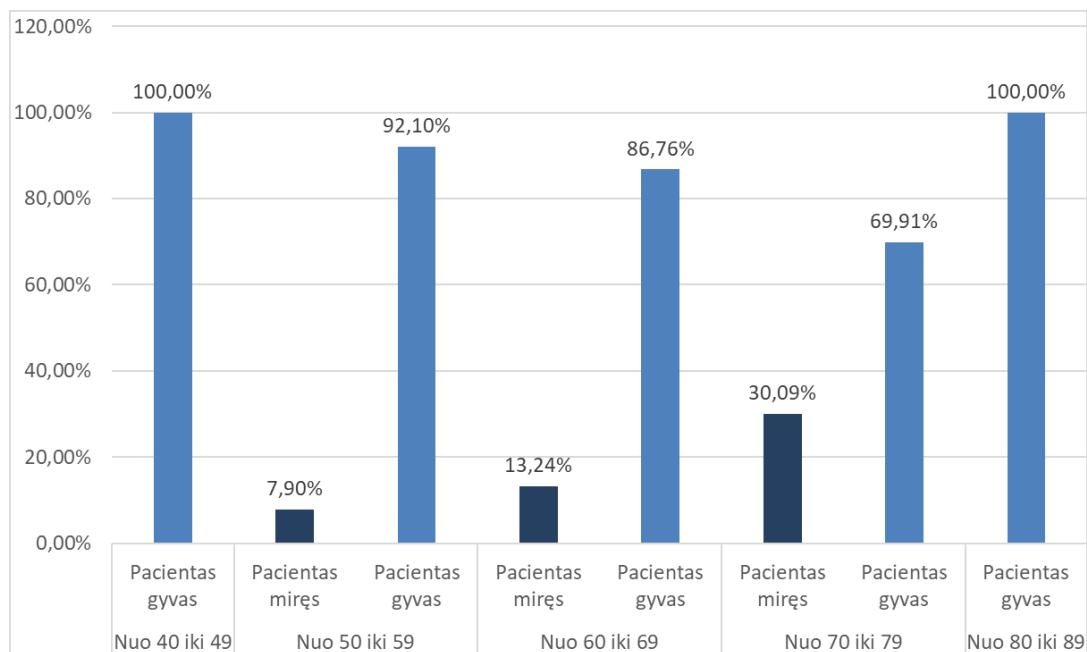
aukščiausias prostatos specifinio antigeno vidurkis buvo amžiaus grupėje nuo 75 iki 79 metų (žr. 19 pav.).

Nors žemiausias PSA rodiklis yra amžiaus grupėje nuo 85 iki 89 metų, tačiau į ją pateko tik vienas pacientas, todėl dėl nagrinėjamo aspekto šis pogrupis laikomas išskirtimi. Taigi žemiausias prostatos specifinio antigeno rodiklio vidurkis yra žmonių nuo 40 iki 44 metų amžiaus grupėje. Jo reikšmė 7,25 (žr. 19 pav.).



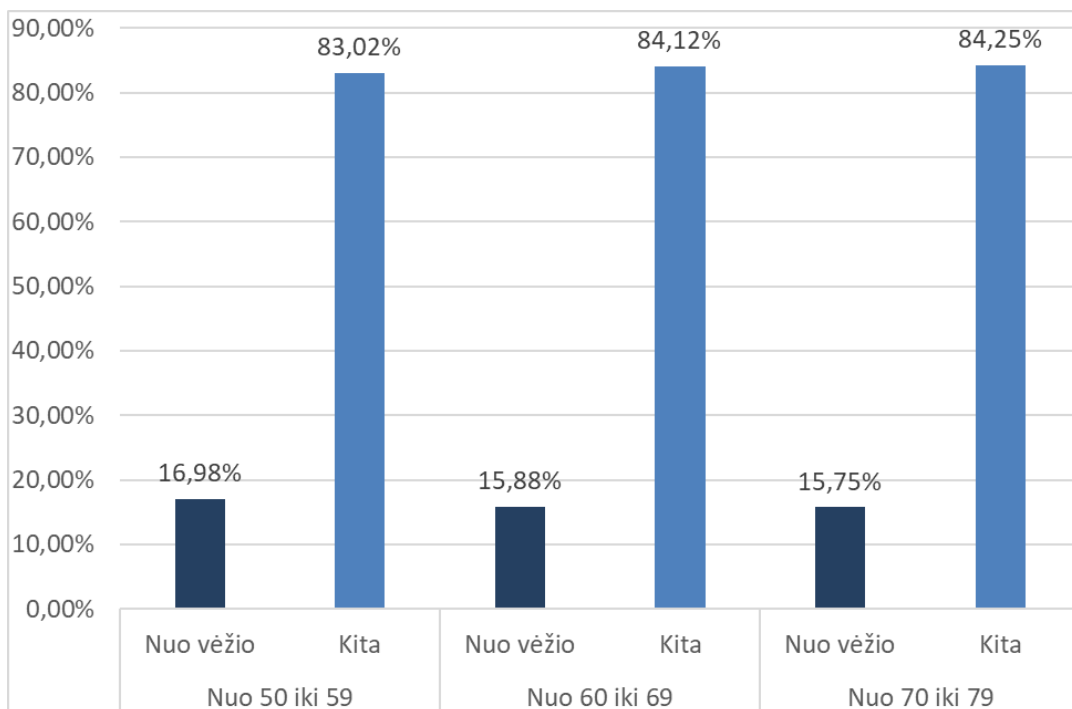
19 pav. Pacientų PSA rodiklio vidurkis amžiaus grupėse

Toliau duomenų imtyje tikrinama paciento būklė, tiksliau, ar pacientas miręs, ar gyvas. Tikrinama atsižvelgiant į amžiaus intervalus kas dešimt metų. Pasitelkus „OS“ kintamąjį, kurio reikšmė yra paciento mirtis dėl bet kurios priežasties, įskaitant ir vėžį, nustatyta, kad duomenų imtyje iš viso yra 2060 gyvų pacientų ir 350 mirusių pacientų (žr. 5 lentelę ir 4 priedą). Tada kiekviena amžiaus grupė laikoma atskira duomenų imtimi. Šis požiūris leidžia palyginti atskiras amžiaus grupes „OS“ kintamojo atžvilgiu. Pastebėta, kad nuo 40 iki 49 metų ir nuo 80 iki 89 metų intervaluose visi pacientai yra gyvi. Taip pat šios imtys yra pačios mažiausios respondentų skaičiaus atžvilgiu, todėl jos nėra labai informatyvios. Grafike matyti, jog pacientui patenkant į vis vyresnę amžiaus grupę (t.y. senstant), procentine išraiška pacientų mirčių skaičius auga (žr. 20 pav.).



20 pav. Pacientų mirties rodiklis amžiaus grupėse, procentais

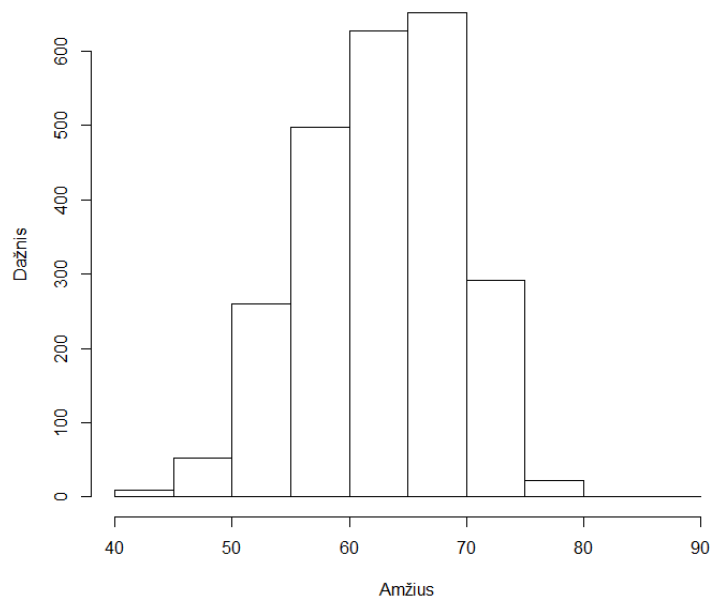
Remiantis ta pačia logika analizuojama tik mirusių pacientų duomenų aibė, siekiant nustatyti mirties priežastis skirtingose amžiaus grupėse. Pasitelkiami kintamieji „CSS“ ir „Death_other_causes“ (žr. 2 lentelę). Iš tyrimo rezultatų galima pastebėti, kad vėžys kaip mirties priežastis diagnozuojama gana retai, tik 16 % – 17 % visų mirties atvejų. Taip pat, nepriklausomai nuo amžiaus grupės, vėžio diagnozė kaip mirties priežastis yra nustatoma panašiam kiekiui pacientų, t.y. stabiliai (žr. 21 pav. ir 5 priedas).



21 pav. Neišgyvenusių skirtingo amžiaus pacientų mirties priežastis, procentais

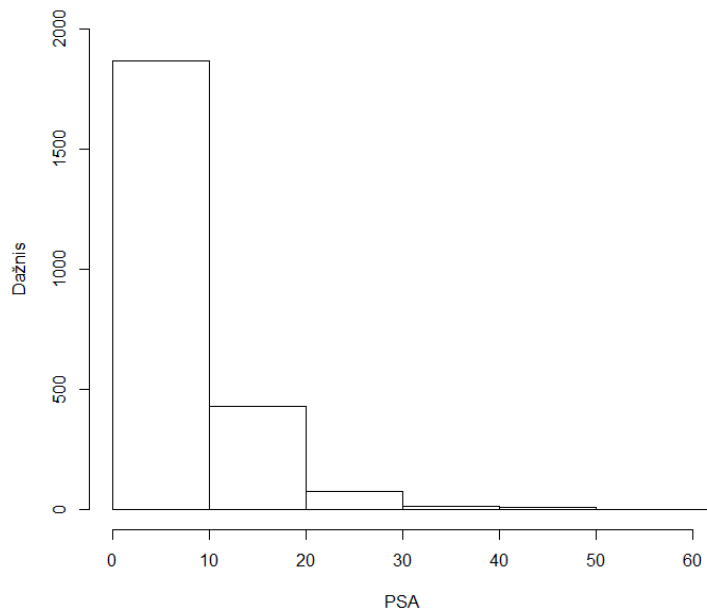
3.3. Koreliacinės analizės modelis

Toliau vykdoma koreliacinė analizė. Patikrinami skaitinių kintamųjų „Amzius“, „PSAI“ ir „SURVIVAL“ pasiskirstymo grafikai (žr. 22, 23 ir 23 pav.). Amžiaus kintamasis turi normalųji skirstinį, grafike matomi stebėjimai yra varpo formos (žr. 22 pav.).



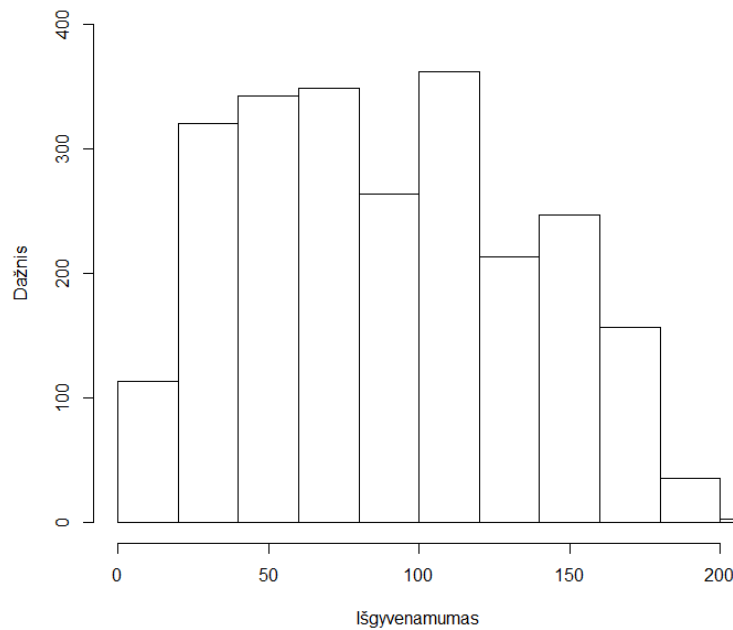
22 pav. Amžiaus pasiskirstymo grafikas

PSA normaliojo pasiskirstymo neturi (žr. 23 pav.).



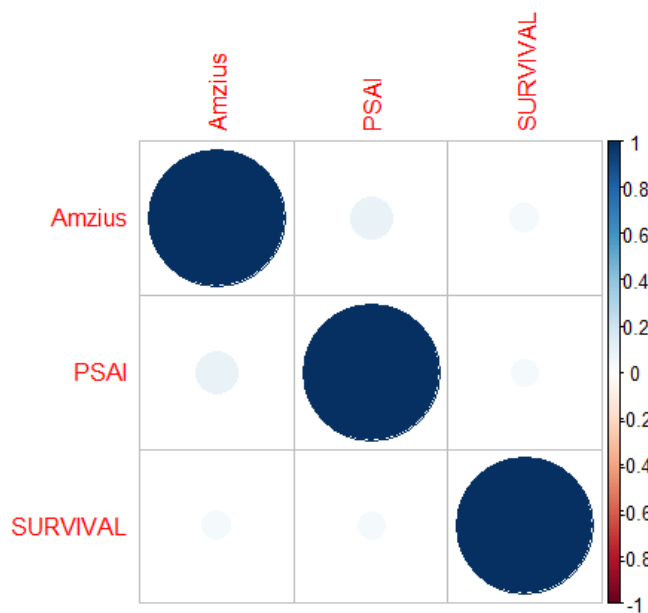
23 pav. PSA pasiskirstymo grafikas

Išgyvenamumo stebėjimai yra maždaug varpo formos (daugiau stebėjimų pasiskirstymo viduryje, mažiau kraštuose) (žr. 24 pav.).



24 pav. Išgyvenamumo pasiskirstymo grafikas

Šiuo atveju Pirsono koreliacija yra negalima, todėl naudojamas Spirmeno koreliacijos metodas. Koreliacijos grafike ir reikšmių lentelėje stiprių ryšių tarp kintamųjų „Amzius“, „PSAI“ ir „SURVIVAL“ nematyti. (žr. 25 pav. ir 6 lentelę)



25 pav. Spirmeno koreliacija trims kintamiesiems

6 lentelė. Spirmeno koreliacijos reikšmės

	Amzius	PSAI	SURVIVAL
Amzius	1,00000000	0,09706619	0,04534989
PSAI	0,09706619	1,00000000	0,04334886
SURVIVAL	0,04534989	0,04334886	1,00000000

3.4. Internetinis puslapis

Interaktyviai sistemai buvo išskirti funkciniai reikalavimai (žr. 7 lentelė.)

7 lentelė. Interaktyvios sistemos funkciniai reikalavimai

Nr.	Reikalavimas
1	Lengvas prieinamumas
2	Paprastas ir aiškus valdymas
3	Pilnas atributų užpildymas
4	Aiškus rezultatų pateikimas

Internetinis puslapis turi ne tik teikti lengvą prieinamumą prie programinio įrankio, bet ir bendrai pateikti mokslo populiariąją informaciją apie priešinės liaukos vėžio ligą, taip pat jos diagnostikos būdus. Kaip šio tikslo rezultatas buvo sukurtas internetinis puslapis, pasiekiamas adresu: <https://indremarijajuskait.wixsite.com/plv-modelis>.

Pagrindiniame, pirmajame skirtuke, pavadintame „PLV“, vartotojas trumpai supažindinamas su šia liga ir jos aktualumu (žr. 6 priedas). Taip pat suteikiama galimybė plačiau susipažinti su tema paspaudus nuorodą „Daugiau informacijos“, kuri informacijos ieškančią asmenį nukreipia į Lietuvos Nacionalinio vėžio instituto internetinį puslapį (žr. 26 pav.).

Daugiau informacijos: <https://www.nvi.lt>

26 pav. Nuoroda į Nacionalinio vėžio instituto internetinį puslapį

PLV skirtukas turi vidinį skirtuką „Diagnostika“ (žr. 27 pav.). Jį atsidaręs vartotojas supažindinamas su bendromis vėžio diagnostikos galimybėmis (žr. 7 ir 8 priedas.)



27 pav. Skirtukas „Diagnostika“

Kitoje skiltyje, pavadintoje „Naudojimosi instrukcija“, supažindinama su žingsniais, kuriuos vartotojas turės atlikti, norėdamas parsisiųsti programą bei ją tinkamai naudoti. Taip pat trumpai supažindinama su tyrimo rezultatais ir jų interpretavimu (žr. 28 pav ir 9, 10 priedai).



28 pav. Skirtukas „Naudojimosi instrukcija“

Skirtuke „Prognozavimo įrankis“ dar kartą primenama, kad prieš parsisiunčiant programą reikia perskaityti skirtuke „Naudojimosi instrukcija“ pateiktą informaciją. Paspaudus mygtuką „Pradėti“, vartotojas nukeliamas į aplanką, kuriame yra saugomas sukurtas programinis įrankis (žr. 29 pav.). Parsisiuntęs failus ir besivadovaujantis instrukcijomis vartotojas gali lengvai naudotis šiuo produktu.



29 pav. Skirtukas „Prognozavimo įrankis“

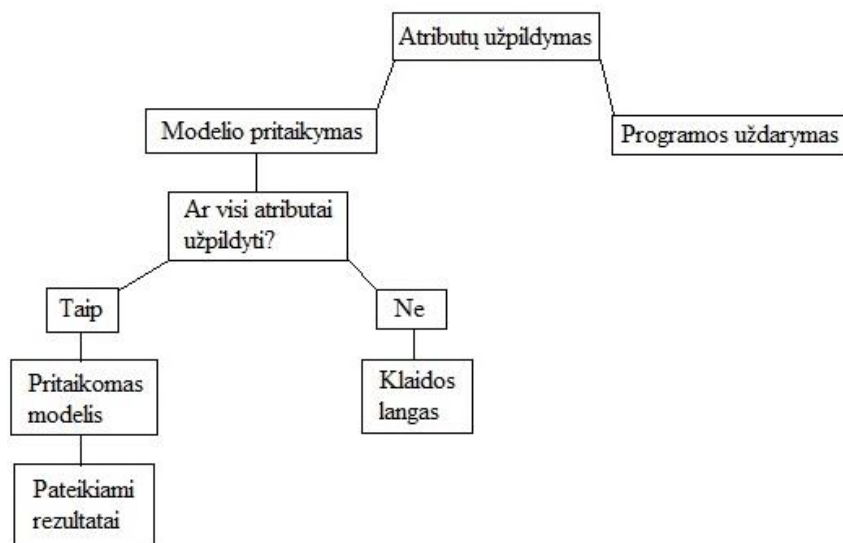
Paskutiniame skirtuke, pavadintame „Kontaktai“, galima rasti trumpą informaciją apie projekto autorę. Taip pat nurodoma autorės kontaktinė informacija ir interaktyvi sąsaja su „Google Maps“ nuoroda į autorės nurodytą adresą (žr. 30 pav.).



30 pav. Skirtukas „Kontaktai“

3.5. Automatizuotas ir statistiniais metodais paremtas modelis, programa

Sukurti šią programą buvo pasitelktas „RStudio“ įrankis su naujo projekto kūrimo galimybe „Shiny Web App“. Šis pasirinkimas suteikė galimybę sukurti programą su interaktyvia sistema ir galimybe lokaliame serveryje (naudojamame tik tame kompiuteryje) atsidaryti programą internetinėje naršyklėje. Projekto interaktyvios sistemos funkcionalumą paaiškina 31 pav.



31 pav. Projekto interaktyvios sistemos funkcionalumas

Įvykdžius programą atsidaro vartotojo grafinė sąsaja. Atsidariusio lango viršuje paspaudus parinktį „Open in Browser“, internetinėje naršyklėje atsidaro toks pats vartotojo grafinis langas (žr. 32 pav.). Pagal patogumą šį įrankį vartotojas gali pasirinkti naudoti tiek grafinėje sąsajoje, tiek internetinėje naršyklėje. Kairėje atsidariusio lango pusėje, pilka spalva atskirtoje skiltyje vartotojas turi užpildyti nurodytas skiltis. Paciento amžius yra skaičiuojamas metais, PSA rodiklis nurodomas iš gautų kraujo tyrimo rezultatų (žr. 32 pav.).

32 pav. Programos „PLV modelis“ vartotojo sąsajos vaizdas

Toliau programa leidžia pasirinkti vieną iš penkių galimų išvešėjusių limfmazgių skaičiaus pasirinkimų (žr. 33 pav.).

33 pav. Programos „PLV modelis“ išvešėjusių limfmazgių skaičiaus pasirinkimai

Vartotojui leidžiama pasirinkti iš trijų patologinės vėžio stadijos galimybių (žr. 34 pav.):

- vėžys nematomas iš tomografijos;
- vėžys formuojasi vienoje vietoje;
- vėžys vienoje vietoje stipriai išsiplėtęs.

PLV modelis

Paciento amžius:

PSA:

Isvesejusių limfmazgių skaičius:
Vienas ▼

Patologinė vėžio stadija:
Vėžys nematomas iš tomografijos ▲
Vėžys nematomas iš tomografijos
Vėžys formuojasi vienoje vietoje
Vėžys vienoje vietoje stipriai išsiplėtęs

Pasalintos prostatos buklė:
Chirurgškai pasalintos prostatos kraštai yra svarūs ▼

Pritaikyti modelį!
Paspauskite mygtuką kuris pritaikys modelį.

Su įvestais rodikliais išgyvenamumas mėnesiais:

Su įvestais rodikliais išgyvenamumas metais:

Išgyventi 1 m. :

Išgyventi 2 m. :

Išgyventi 5 m. :

34 pav. Programos „PLV modelis“ patologinės vėžio stadijos pasirinkimai

Kitas žingsnis – nurodoma limfmazgių būklė (žr. 35 pav.):

- limfmazgiai švarūs;
- limfmazgiai pažeisti vėžiu;
- limfmazgiai nešalinti, būklė nežinoma.

PLV modelis

Paciento amžius:

PSA:

Isvesejusių limfmazgių skaičius:
Vienas ▼

Patologinė vėžio stadija:
Vėžys nematomas iš tomografijos ▼

Limfmazgių buklė:
Limfmazgiai svarūs ▲
Limfmazgiai svarūs
Limfmazgiai pažeisti vėžiu
Limfmazgiai nešalinti, buklė nežinoma

Pritaikyti modelį!
Paspauskite mygtuką kuris pritaikys modelį.

Su įvestais rodikliais išgyvenamumas mėnesiais:

Su įvestais rodikliais išgyvenamumas metais:

Išgyventi 1 m. :

Išgyventi 2 m. :

Išgyventi 5 m. :

35 pav. Programos „PLV modelis“ limfmazgių būklės pasirinkimai

Paskutinis pasirinkimas leidžia vartotojui pasirinkti pašalintos prostatos būklę iš duotų trijų galimybių (žr. 36 pav.):

- chirurgiškai pašalintos prostatos kraštai yra švarūs;
- chirurgiškai pašalintos prostatos kraštai yra su vėžinėmis ląstelėmis;
- nėra aišku.

PLV modelis

Paciento amžius:

PSA:

Isvesejusių limfmazgių skaičius:
Vienas ▾

Patologinė vezio stadija:
Vezys nematomas iš tomografijos ▾

Limfmazgių bukle:
Limfmazgiai svarūs ▾

Pasalintos prostatos bukle:
Chirurgiškai pašalintos prostatos kraštai yra svarūs ▲

Chirurgiškai pašalintos prostatos kraštai yra svarūs

Chirurgiškai pašalintos prostatos kraštai yra su vėžinėmis ląstelėmis

Nėra aišku

Su įvestais rodikliais išgyvenamumas mėnesiais:

Su įvestais rodikliais išgyvenamumas metais:

Išgyventi 1 m. :

Išgyventi 2 m. :

Išgyventi 5 m. :

36 pav. Programos „PLV modelis“ pašalintos prostatos būklės pasirinkimui

Paspaudus mygtuką „Pritaikyti modelį!“ programa pritaiko modelį ir suskaičiuoja rezultatus. Rezultatai pateikiami programinio lango dešinėje pusėje.

Visa turima duomenų aibė padalinama į dvi dalis. 80 % duomenų 1928 eilutės paskiriamos apmokymui, o likusieji 20 % duomenų 482 eilutės testavimui. Testavimo duomenų aprašomojoje statistikoje „SURVIVAL“ atributo minimumas – 1, maksimumas – 73 (žr. 37 pav.).

Amžius	PSAI	GLpoRP_ISUP	pT	LNI01	R1
Min. :40.00	Min. : 1.890	1: 43	0:245	0:153	0:317
1st Qu.:57.00	1st Qu.: 4.662	2:257	1:180	1: 34	1:162
Median :62.00	Median : 6.240	3: 88	2: 57	2:295	2: 3
Mean :61.97	Mean : 8.590	4: 27			
3rd Qu.:67.00	3rd Qu.: 9.600	5: 67			
Max. :79.00	Max. :82.000				
SURVIVAL	OS	CSS	Death_other_causes	Age64vs65	
Min. : 1.00	0:468	0:479	0:471	0:293	
1st Qu.:25.00	1: 14	1: 3	1: 11	1:189	
Median :34.00					
Mean :34.08					
3rd Qu.:42.75					
Max. :73.00					

37 pav. Aprašomosios statistikos rezultatų vizualizacija R programinėje įrangoje testavimo duomenims

Taip pat testavimo duomenyse, pagal „SURVIVAL“ atributą, patikrinama koks yra realus respondentų išgyvenamumas analizuojamais laikotarpiais: 1 metai, 2 metai ir 5 metai (žr. 8 lentelę).

8 lentelė. Testavimo duomenų „SURVIVAL“ atributas

Intervalas „SURVIVAL“ reikšmei	Respondentų skaičius
[1;11], iki 1 metų	11
[11;73], daugiau nei 1 metai	471
[1;23], iki 2 metų	100
[24;73], daugiau nei 1 metai	382
[1;75], iki 5 metų	482
daugiau nei 5 metai	0

Toliau ieškomas geriausias dirbtinių neuroninių tinklų modelis vieneriems, dvejiems ir penkeriems metams. Kiekvienam iš šių trijų modelių testuojami keturi skirtingi dirbtiniai neuroniniai tinklai. Jie pasirenkami atsitiktinai. Toliau, naudojant testinius duomenis, kiekvienam bandymui skaičiuojama susimaišymo matrica ir apskaičiuojamas modelio tikslumas, taip pat parodoma kiekvieno mokymo proceso žingsnių skaičius (žr. 9, 10 ir 11 lentelės).

Atlikus testavimus pasirinkti šie modeliai:

1 metams (žr. 9 lentelę):

- dviejų užslėptų sluoksnių;
- pirmam sluoksnyje 4, o antram – 3 neuronai;
- aukščiausias modelio prognozės tikslumas 97 %.

9 lentelė. 1 metų modelio pasirinkimas

Užsklėptų sluoksnių sk.	Neuronų sk.	Susimaišymo matrica				Modelio prognozės tikslumas %	Mokymosi žingsnių sk.
		Teisingai neigiami 0/0	Klaidingai neigiami 0/1	Klaidingai teigiami 0/1	Teisingai teigiami 1/1		
1	4	0	11	20	451	94%	5655
2	1	1	10	15	456	95%	825
	3						
2	4	2	9	7	464	97%	23779
	3						
3	3	0	11	6	465	96%	21599
	2						
	3						

2 metams (žr. 10 lentelę):

- dviejų užslėptų sluoksnių;
- pirmam sluoksnyje 3, o antram – 1 neuronai;
- aukščiausias modelio prognozės 83 %.

10 lentelė. 2 metų modelio pasirinkimas

Užsklėptų sliuksnių sk.	Neuronų sk.	Susimaišymo matrica				Modelio prognozės tikslumas %	Mokymosi žingsnių sk.
		Teisingai neigiami 0/0	Klaidingai neigiami 0/1	Klaidingai teigiami 0/1	Teisingai teigiami 1/1		
1	5	1	99	9	373	78%	4358
2	2	1	99	9	373	78%	20126
	2						
2	3	27	73	7	375	83%	28652
	1						
3	2	0	100	0	382	79%	7215
	2						
	1						

5 metams (žr. 11 lentelę):

- dviejų užslėptų sluoksnių;
- pirmam sluoksnyje 1, o antram – 2 neuronai;
- aukščiausias modelio prognozės 65 %.

11 lentelė. 5 metų modelio pasirinkimas

Užsklėptų sliuksnių sk.	Neuronų sk.	Susimaišymo matrica				Modelio prognozės tikslumas %	Mokymosi žingsnių sk.
		Teisingai neigiami 0/0	Klaidingai neigiami 0/1	Klaidingai teigiami 1/0	Teisingai teigiami 1/1		
1	3	31	451	0	0	6%	21028
2	1	311	171	0	0	65%	33476
	2						
2	1	193	289	0	0	40%	7512
	3						
3	2	218	264	0	0	45%	4900
	1						
	2						

Atliekant testavimą pastebėta, kad didesnis neuronų sluoksnių skaičius nebūtinai apskaičiavo geresnius rezultatus. Testavimo metu ieškomas balansas tarp užslėptų sluoksnių ir juose esančių neuronų skaičiaus.

Pasirinkti geriausi modeliai pritaikomi programoje. Paspaudus mygtuką „Pritaikyti modelį!“ programa pritaiko modelį ir suskaičiuoja rezultatus. Nurodytu konkrečiu atveju, kai pacientui yra 65 metai, jo PSA rodiklis yra 2, išvešėjusių limfmazgių skaičius yra trys, diagnostinėje skiltyje apie patologinę vėžio stadiją nurodoma, kad vėžys formuojasi vienoje vietoje, limfmazgiai pažeisti vėžio, o dėl prostatos būklės nustatyta, kad chirurgiškai pašalintos prostatos kraštai yra su vėžinėmis ląstelėmis. Šiuo atveju rezultatų skiltyje matomi apskaičiuoti galimybės išgyventi rodikliai. Pirmasis rezultatas nurodo, kiek tikėtina pacientas išgyvens skaičiuojant mėnesiais (žr. 38 pav.). Antrasis rezultatas suskaičiuoja pirmąjį rezultatą metais. Dėl analizuojamo konkretaus pavyzdžio tikėtina, kad

pacientas išgyvens šiek tiek daugiau nei šešerius metus. Pasitelkus pasirinktus geriausius dirbtinius neuroninius tinklus, apskaičiuotas rezultatas, kad pacientas, turintis būtent šiuos rodiklius, išgyvens vienerius metus (žr. 38 ir 39 pav.). Kuo arčiau vieneto yra gautas rezultatas (užrašytas žemiau \$net.result), tuo labiau tikėtina, kad pacientas išgyvens 1 metus. Taip pat rezultatų lange išspausdinami kiekvieno neurono tarpiniai rezultatai.

PLV modelis

Paciento amžius:

PSA:

Išvėjęs lymfagzgiu skaičius:

Patologine vezio stadija:

Limfmazgiu bukle:

Pasalintos prostatos bukle:

Paspauskite mygtuka kuris pritaikys modeli.

Su įvestais rodikliais išgyvenamumas mėnesiais:

76.61391

Su įvestais rodikliais išgyvenamumas metais:

6.384493

Tikimybė išgyventi 1 m. :

```
$neurons
$neurons[[1]]
  Amzius PSAI GLpoRP_ISUP pT LNI01 R1
[1,] 1    65    2            3 1    1 1

$neurons[[2]]
  [,1] [,2] [,3] [,4] [,5]
[1,] 1 6.475741e-05 1 3.130954e-28 1.659061e-09

$neurons[[3]]
  [,1] [,2] [,3] [,4]
[1,] 1 0.7076528 0.7231947 0.7999545

$net.result
  [,1]
[1,] 0.988555
```

Tikimybė išgyventi 2 m. :

```
$neurons
$neurons[[1]]
  Amzius PSAI GLpoRP_ISUP pT LNI01 R1
[1,] 1    65    2            3 1    1 1

$neurons[[2]]
  [,1] [,2] [,3] [,4]
[1,] 1 0.9995289 1 1

$neurons[[3]]
  [,1] [,2]
[1,] 1 1

$net.result
  [,1]
[1,] 0.9995083
```

Tikimybė išgyventi 5 m. :

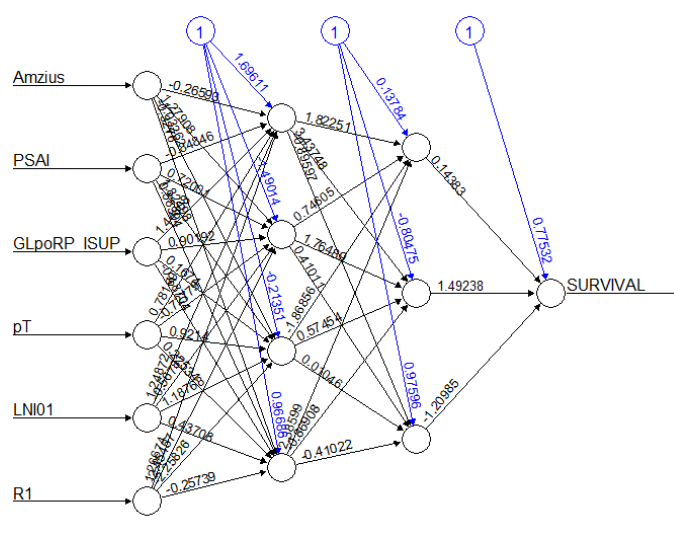
```
$neurons
$neurons[[1]]
  Amzius PSAI GLpoRP_ISUP pT LNI01 R1
[1,] 1    65    2            3 1    1 1

$neurons[[2]]
  [,1] [,2]
[1,] 1 1

$neurons[[3]]
  [,1] [,2] [,3]
[1,] 1 0.7486192 0.5101816

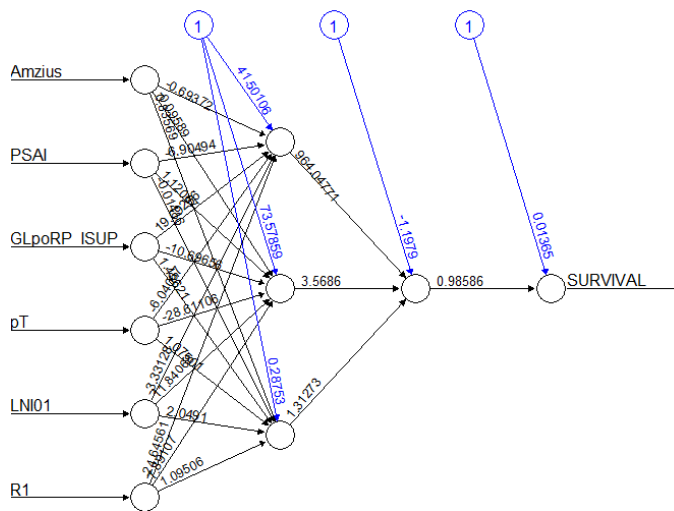
$net.result
  [,1]
[1,] 0.8482282
```

38 pav. Programos įvesčių ir rezultatų vaizdas



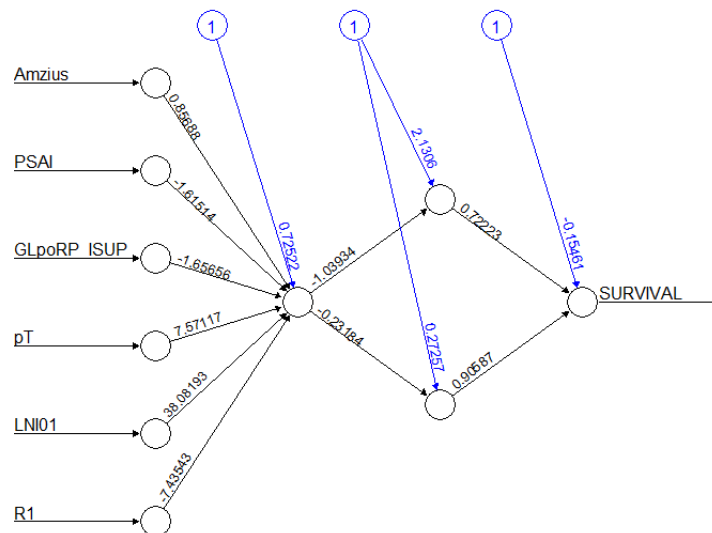
39 pav. Neuroninis tinklas 1 metams

Kitas pasirinktas modelis apskaičiuoja 2 metų išgyvenamumą. Pirmasis lygis turi tris neuronus, o antrasis turi vieną neuroną. Pažvelgus į gautus rezultatus su nurodytomis įvestimis, apskaičiuota, kad pacientas išgyvens dar dvejus metus (žr. 38 ir 40 pav.).



40 pav. Neuroninis tinklas 2 metams

Dar kitas pasirinktas modelis skaičiuoja galimybę išgyventi 5 metus, apskaičiuojamas naudojant taip pat du užslėptus lygius su pirmuoju turinčiu vieną neuroną, o antruoju turinčiu du neuronus. Pažvelgus į gautus rezultatus išgyvenamumas su nurodytomis įvestimis penkiems metams taip pat yra tikėtinas (žr. 38 ir 41 pav.).



41 pav. Neuroninis tinklas 5 metams

Bendrai palyginus gautus rezultatus matyti, jog tikėtinas pacientui išgyventi daugiau metų nuosekliai mažėja. Regresijos būdu apskaičiuota gyvenimo trukmė buvo apie šešeri metai, o ir tolimesnė neuroninių tinklų rezultatų interpretacija yra tinkama, nes tikėtinas išgyventi, didėjant metų skaičiui, mažėja.

Išvados ir rekomendacijos

Magistro darbo metu visi išsikelti uždaviniai buvo atlikti, o užsibrėžtas tikslas – pasiektas.

1. Išanalizavus pasaulyje atliktus tyrimus nustatyta, kad priešinės liaukos vėžio diagnozė yra ypač svarbi ne tik Lietuvos, bet ir viso pasaulio mastu. Lietuvoje prostatos vėžys yra dažniausia vyrų onkologinė liga, kurios sergamumas kas metus vis didėja. Be to, dėl spartaus žmonių populiacijos augimo tikėtina, kad ateityje prostatos vėžio atvejų skaičius taip pat išaugs.
2. Paruoštas magistro baigiamojo projekto duomenų rinkinys analizei.
3. Sukurtas projekto tyrimo modelis.
4. Duomenų imtyje pastebėta, kad vėžys kaip mirties priežastis diagnozuojama retai, 16 % – 17 % visų mirties atvejų. Dažniau diagnozuojamos kitos mirties priežastys.
5. Sukurtas lengvai prieinamas internetinis puslapis su programos instaliavimo instrukcija, supažindinimu su priešinės liaukos vėžiu bei jos diagnostikos būdais ir nuoroda į pačią programą. Jis pasiekiamas adresu: <https://indremarijajuskait.wixsite.com/plv-modelis>.
6. Panaudojus sudarytą modelį ir programines priemones sukurta interaktyvi sistema. Ši programa lengvai prieinama vartotojams.
7. Sukurtas galutinis produktas prognozuoja, pagal visus pilnai nurodytus atributus, pacientų išgyvenamumą. Išsiaiškinta, kad 1 metų išgyvenamumo modelis prognozuoja 97 %, 2 metų modelis 83 % ir 5 metų modelis 65 % tikslumu. Pastebėta, kad pasirenkant dirbtinius neuroninius tinklus rezultato tikslumas priklausys nuo tinkamai parinklų užslėptų sluoksnių ir juose esančių neuronų kiekį.

Rekomendacijos:

- Programą naudoti gydytojams urologams prognozuojant paciento išgyvenamumo laikotarpį, pagal kurį būtų parenkamas geriausias gydymo motodas.
- Surinkti didesnę duomenų aibę kuria apmokyti duomenis, taip ateityje būtų galima modelius prognozuoti su dar aukštesniu tikslumu.

Literatūros sąrašas

1. American Cancer Society. [žiūrėta 2021-03-01] Prieiga per internetą: <https://www.cancer.org>.
2. National Cancer Institute. [žiūrėta 2021-03-01] Prieiga per internetą: <https://www.cancer.gov/publications/dictionaries/cancer-terms/def/cancer>.
3. Nacionalinis vėžio institutas. Onkologinės ligos. (2021-01-20) [žiūrėta 2021-03-01] Prieiga per internetą: <https://www.nvi.lt/onkologines-ligos/>.
4. Cancer cell. illustration showing cancer disease development [žiūrėta 2021-03-02] Prieiga per internetą: <https://www.dreamstime.com/cancer-cell-normal-healthy-tissue-malignant-tumour-illustration-showing-disease-development-vector-diagram-your-design-image135725846>.
5. Nacionalinis vėžio institutas. Statistiniai duomenys. (2020-11-20) [žiūrėta 2021-03-02] Prieiga per internetą: <https://www.nvi.lt/naujausi-duomenys/>.
6. Ferlay J, Colombet M, Soerjomataram I, et al. Cancer statistics for the year 2020: An overview. Int. J. Cancer. 2021;1–12. [žiūrėta 2021-03-02] doi: 10.1002/ijc.33588 Prieiga per internetą: https://vb.ktu.edu/primo-explore/fulldisplay?docid=TN_cdi_proquest_miscellaneous_2508888045&context=PC&vid=KTU&lang=lt-LT&search_scope=KTU&adaptor=primo_central_multiple_fe&tab=default_tab&query=any,contains,Cancer%20statistics%20for%20the%20year%202020:%20An%20overview&offset=0.
7. World Population. [žiūrėta 2021-03-02] Prieiga per internetą: <https://www.worldometers.info/world-population/>.
8. Nacionalinis vėžio institutas. Priešinės liaukos vėžys. (2021-01-20) [žiūrėta 2021-03-02] Prieiga per internetą: <https://www.nvi.lt/prostatos/>.
9. Medical Definition of DNA (2021-03-29) [žiūrėta 2021-03-31] Prieiga per internetą: <https://www.medicinenet.com/dna/definition.htm>.
10. What is DNA and how does it work? [žiūrėta 2021-03-05] Prieiga per internetą: <https://www.medicalnewstoday.com/articles/319818>.
11. PRIEŠINĖS LIAUKOS VĖŽIO ANKSTYVOSIOS DIAGNOSTIKOS FINANSAVIMO PROGRAMA. Patvirtinta Lietuvos Respublikos sveikatos apsaugos ministro 2005 m. gruodžio 14 d. įsakymu Nr. V-973. Prieiga per internetą: https://e-seimas.lrs.lt/portal/legalAct/lt/TAD/TAIS.268622?jfwid=bkaxlm28&fbclid=IwAR0tnFcAoAbtbgRehA5bC3I8Ouxna_UgbK94HQbp50qwDVt6DpVmjtTitFBc.
12. Borley N., R Fenely M., Prostate cancer: diagnosis and staging. Asian J Androl. 2009; 11(1): 74–80. [žiūrėta 2021-05-21]. doi: 10.1038/aja.2008.19
13. American Cancer Society. Prostate Cancer Stages and Other Ways to Assess Risk. [žiūrėta 2021-03-06] Prieiga per internetą: <https://www.cancer.org/cancer/prostate-cancer/detection-diagnosis-staging/staging.html>.
14. Four stages of prostate cancer. [žiūrėta 2021-03-07] Prieiga per internetą: <https://www.dreamstime.com/four-stages-prostate-cancer-malignant-tumor-spreading-image109383162>.
15. Internet Definition. [žiūrėta 2021-03-15] Prieiga per internetą: <https://techterms.com/definition/internet>.

16. What is the difference between webpage, website, web server, and search engine? [žiūrėta 2021-03-15] Prieiga per internetą: https://developer.mozilla.org/en-US/docs/Learn/Common_questions/Pages_sites_servers_and_search_engines.
17. The 21 Best Free Website Builders: We Expose Their Strengths and Weaknesses. (2021 03 10) [žiūrėta 2021-03-15] Prieiga per internetą: <https://www.websitetooltester.com/en/blog/best-free-website-builders/>.
18. WIX [žiūrėta 2021-03-05] Prieiga per internetą: <https://www.wix.com/about/us>.
19. Introduction to R. [žiūrėta 2021-03-10] Prieiga per internetą: <https://www.r-project.org/about.html>.
20. What is R Programming Language? Introduction & Basics. [žiūrėta 2021-03-11] Prieiga per internetą: <https://www.guru99.com/r-programming-introduction-basics.html>.
21. Eglė Butkevičienė ir Aida Vaicekauskaitė. KIEKYBINIŲ DUOMENŲ INTERNETINIŲ ARCHYVŲ ANALIZĖ [žiūrėta 2021-03-13] Prieiga per internetą: http://www.lidata.eu/index.php?file=files/mokymai/kiek1/kiek1.html&course_file=kiek1_3_2.html.
22. Visuotinė lietuvių enciklopedija. [žiūrėta 2021-03-14] Prieiga per internetą: <https://www.vle.lt/rodykle/>.
23. Skaitinės suvestinės. [žiūrėta 2021-03-15] Prieiga per internetą: <https://mokymai.github.io/biostatistika/skaitines-suvesines.html>.
24. Vencloviene J. Statistiniai metodai medicinoje. Kaunas: Vytauto didžiojo universitetas (2010). [žiūrėta 2021-03-20] Prieiga per internetą: <https://www.vdu.lt/cris/bitstream/20.500.12259/271/1/ISBN9789955125587.pdf>.
25. Gogtay NJ., Thatte UM. Principles of Correlation Analysis. Journal of The Association of Physicians of India. 2017 Kovas; 65(3):78-81. PMID: 28462548. [žiūrėta 2021-04-05] Prieiga per internetą: https://www.kem.edu/wp-content/uploads/2012/06/9-Principles_of_correlation-1.pdf.
26. Correlation (Pearson, Kendall, Spearman). [žiūrėta 2021-04-10] Prieiga per internetą: <https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/correlation-pearson-kendall-spearman/#:~:text=Correlation%20is%20a%20bivariate%20analysis,between%20%2B1%20and%20%2D1>.
27. Koreliacinė analizė: ryšys tarp kintamųjų. [žiūrėta 2021-04-11] Prieiga per internetą: <https://mokymai.github.io/biostatistika/koreliacija.html>.
28. Rene Y. Choi, Aaron S. Coyner, Jayashree Kalpathy-Cramer, Michael F. Chiang, J. Peter Campbell. Introduction to Machine Learning, Neural Networks, and Deep Learning. Translational Vision Science & Technology February 2020, Vol. 9, 14. doi:<https://doi.org/10.1167/tvst.9.2.14> [žiūrėta 2021-03-15] Prieiga per internetą: <https://tvst.arvojournals.org/article.aspx?articleid=2762344>.
29. Andrius Berniukevičius, Jevgenij Kurilov. Dirbtiniai neuroniniai tinklai personalizuotam mokymuisi. Lietuvos matematikos rinkinys. Ser. B. Vilnius : Matematikos ir informatikos institutas. 2017, T. 58, p. 22-27. ISSN 0132-2818. eISSN 2335-898X. [žiūrėta 2021-03-18] Prieiga per internetą: <https://epublications.vu.lt/object/elaba:25739677/>.

30. Wollseifen T. Using R to build artificial neural networks in medical data. 2019. [žiūrėta 2021-04-25] Prieiga per internetą: <https://www.lexjansen.com/phuse/2019/ml/ML06.pdf>.
31. Markūnaitė R. Neuroniniai tinklai. 2006. [žiūrėta 2021-04-26] Prieiga per internetą: <http://www.elektronika.lt/teorija/kompiuterija/4342/neuroniniai-tinklai/>.
32. Ekonometriniai modeliai. [žiūrėta 2021-04-27] Prieiga per internetą: http://www.ilab.lt/stabingiene/sk2_2.html.
33. Regresinė analizė. [žiūrėta 2021-05-01] Prieiga per internetą: <https://mokymai.github.io/biostatistika/tiesine-regresija.html>.
34. Janilionis V., DAUGIALYPĖS REGRESINĖS ANALIZĖS TAIKYMAS SOCIALINIUOSE TYRIMUOSE. [žiūrėta 2021-05-02] Prieiga per internetą: http://www.lidata.eu/index.php?file=files/mokymai/Janilionis_IV/jan_IV.html&course_file=jan_IV_3.html.

Priedai

1 priedas. Diagnozuotų onkologinių susirgimų skaičius ir sergamumo rodikliai Lietuvoje 2015 metais. Vyrų

Lokalizacijos	TLK-10 kodas	Atvejai
Piktybiniai navikai	C00-C96	9319
Lūpos	C00	11
Burnos ertmės ir ryklės	C01-C14	292
Stemplės	C15	180
Skrandžio	C16	491
Gaubtinės žarnos	C18	421
Tiesiosios žarnos, išangės	C19-C21	343
Kepenų	C22	142
Tulžies pūslės, ekstrahepatinių takų	C23, C24	37
Kasos	C25	244
Kitų virškinimo sistemos organų	C17, C26, C48	30
Nosies ertmės, vid. ausies ir ančių	C30, C31	7
Gerklų	C32	163
Plaučių, trachėjos, bronchų	C33, C34	1191
Kitų kvėpavimo sistemos organų	C37-C39	12
Kaulų ir jungiamojo audinio	C40-C41, C45-C47, C49	53
Odos melanoma	C43	139
Kiti odos piktybiniai navikai	C44	823
Krūties	C50	8
Priešinės liaukos	C61	2967
Sėklidžių	C62	29
Kitų lyties organų	C60, C63	23
Inkstų	C64	436
Šlapimo pūslės	C67	300
Kitų šlapimą išskiriančių organų	C65, C66, C68	18
Akių	C69	13
Smegenų	C70-C72	129
Skydliaukės	C73	62
Kitų endokrininių liaukų	C74-C75	13
Nepatikslintos lokalizacijos	C76-C80	227
Hodžkino limfomos	C81	18
Ne Hodžkino limfomos	C82-C85	165
Mielominės ligos	C90	78
Leukemijos	C91-C95	249
Kiti limfinio, kraujodaros audinių	C88, C96	5
Melanoma in situ	D03	13

Krūties navikai in situ	D05	0
Šlapimo pūslės in situ	D09.0	108
Nervų sistemos gerybiniai navikai	D32, D33	55
Kiti nervų sistemos	D42, D43	19
Limfinio ir kraujodaros audinių	D45-D47	171

2 priedas. Pagrindiniai vėžio tipai pagal naujus atvejus (sergamumą) ir mirtis (mirtingumą) pagal lytį kiekviename iš 20 pasaulio regionų 2020 m. [6]

	Male						Female					
	Incidence			Mortality			Incidence			Mortality		
	First	Second	Third	First	Second	Third	First	Second	Third	First	Second	Third
World	Lung	Prostate	Non-melanoma skin	Lung	Liver	Stomach	Breast	Lung	Cervix uteri	Breast	Lung	Cervix uteri
Africa	Prostate	Liver	Lung	Prostate	Liver	Lung	Breast	Cervix uteri	Liver	Breast	Cervix uteri	Liver
Eastern Africa	Prostate	Kaposi sarcoma	NHL	Prostate	Oesophagus	Liver	Cervix uteri	Breast	Oesophagus	Cervix uteri	Breast	Oesophagus
Middle Africa	Prostate	Liver	NHL	Prostate	Liver	NHL	Breast	Cervix uteri	NHL	Cervix uteri	Breast	Liver
Northern Africa	Liver	Lung	Prostate	Liver	Lung	Bladder	Breast	Liver	Cervix uteri	Breast	Liver	Ovary
Southern Africa	Prostate	Lung	Non-melanoma skin	Lung	Prostate	Oesophagus	Breast	Cervix uteri	Non-melanoma skin	Cervix uteri	Breast	Lung
Western Africa	Prostate	Liver	NHL	Prostate	Liver	NHL	Breast	Cervix uteri	Ovary	Breast	Cervix uteri	Liver
Americas	Prostate	Non-melanoma skin	Lung	Lung	Prostate	Colon	Breast	Non-melanoma skin	Lung	Lung	Breast	Colon
Northern America	Non-melanoma skin	Prostate	Lung	Lung	Prostate	Pancreas	Breast	Non-melanoma skin	Lung	Lung	Breast	Pancreas
Caribbean Central America	Prostate	Lung	Colon	Prostate	Lung	Colon	Breast	Lung	Lung	Breast	Lung	Colon
South America	Prostate	Stomach	Colon	Prostate	Stomach	Liver	Breast	Cervix uteri	Thyroid	Breast	Cervix uteri	Liver
Asia	Lung	Stomach	Colon	Lung	Prostate	Stomach	Breast	Cervix uteri	Thyroid	Breast	Lung	Cervix uteri
Eastern Asia	Lung	Stomach	Liver	Lung	Liver	Stomach	Breast	Lung	Cervix uteri	Lung	Breast	Cervix uteri
South-Eastern Asia	Lung	Liver	Prostate	Lung	Liver	Stomach	Breast	Cervix uteri	Colon	Breast	Breast	Stomach
South-Central Asia	Lip and oral cavity	Lung	Stomach	Lung	Lip and oral cavity	Oesophagus	Breast	Cervix uteri	Ovary	Breast	Cervix uteri	Ovary
Western Asia	Lung	Prostate	Bladder	Lung	Stomach	Prostate	Breast	Thyroid	Lung	Breast	Lung	Stomach

(Continues)

	Male						Female					
	Incidence			Mortality			Incidence			Mortality		
	First	Second	Third	First	Second	Third	First	Second	Third	First	Second	Third
Europe	Prostate	Lung	Non-melanoma skin	Lung	Prostate	Colon	Breast	Lung	Colon	Breast	Lung	Colon
Eastern Europe	Lung	Prostate	Colon	Lung	Prostate	Stomach	Breast	Corpus uteri	Colon	Breast	Lung	Colon
Northern Europe	Prostate	Non-melanoma skin	Lung	Lung	Prostate	Colon	Breast	Lung	Colon	Lung	Breast	Colon
Southern Europe	Prostate	Lung	Bladder	Lung	Colon	Prostate	Breast	Colon	Lung	Breast	Lung	Colon
Western Europe	Prostate	Non-melanoma skin	Lung	Lung	Prostate	Colon	Breast	Non-melanoma skin	Lung	Breast	Lung	Pancreas
Oceania	Non-melanoma skin	Prostate	Melanoma of skin	Lung	Prostate	Colon	Non-melanoma skin	Breast	Melanoma of skin	Lung	Breast	Colon
Australia/New Zealand	Non-melanoma skin	Prostate	Melanoma of skin	Lung	Prostate	Colon	Non-melanoma skin	Breast	Melanoma of skin	Lung	Breast	Colon
Melanesia	Prostate	Lip and oral cavity	Lung	Liver	Lung	Prostate	Breast	Cervix uteri	Thyroid	Breast	Cervix uteri	Liver
Micronesia/Polynesia	Prostate	Lung	Liver	Lung	Prostate	Liver	Breast	Lung	Thyroid	Lung	Breast	Ovary

3 priedas Aprašomosios statistikos rezultatų vizualizacija R programinėje įrangoje

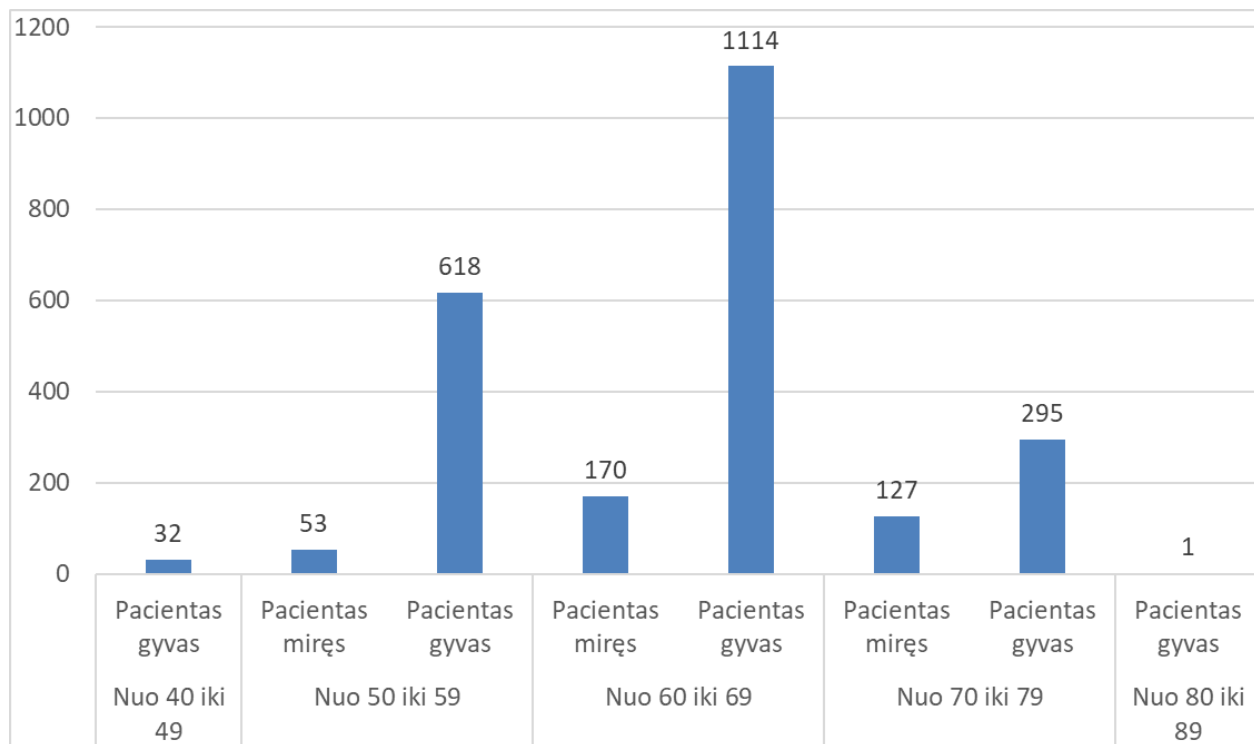
```

      Nr              Amzius              PSAI              GLpoRP_ISUP pT              LNI01
Min.   : 1.0          Min.   :40.00          Min.   : 0.440          1: 630          0:1565          0: 710
1st Qu.: 603.2        1st Qu.:59.00          1st Qu.: 4.700          2:1283          1: 661          1:  83
Median :1205.5        Median :64.00          Median : 6.300          3: 250          2: 184          2:1617
Mean   :1205.5        Mean   :63.19          Mean   : 8.254          4: 109
3rd Qu.:1807.8        3rd Qu.:68.00          3rd Qu.: 9.467          5: 138
Max.   :2410.0        Max.   :87.00          Max.   :98.400

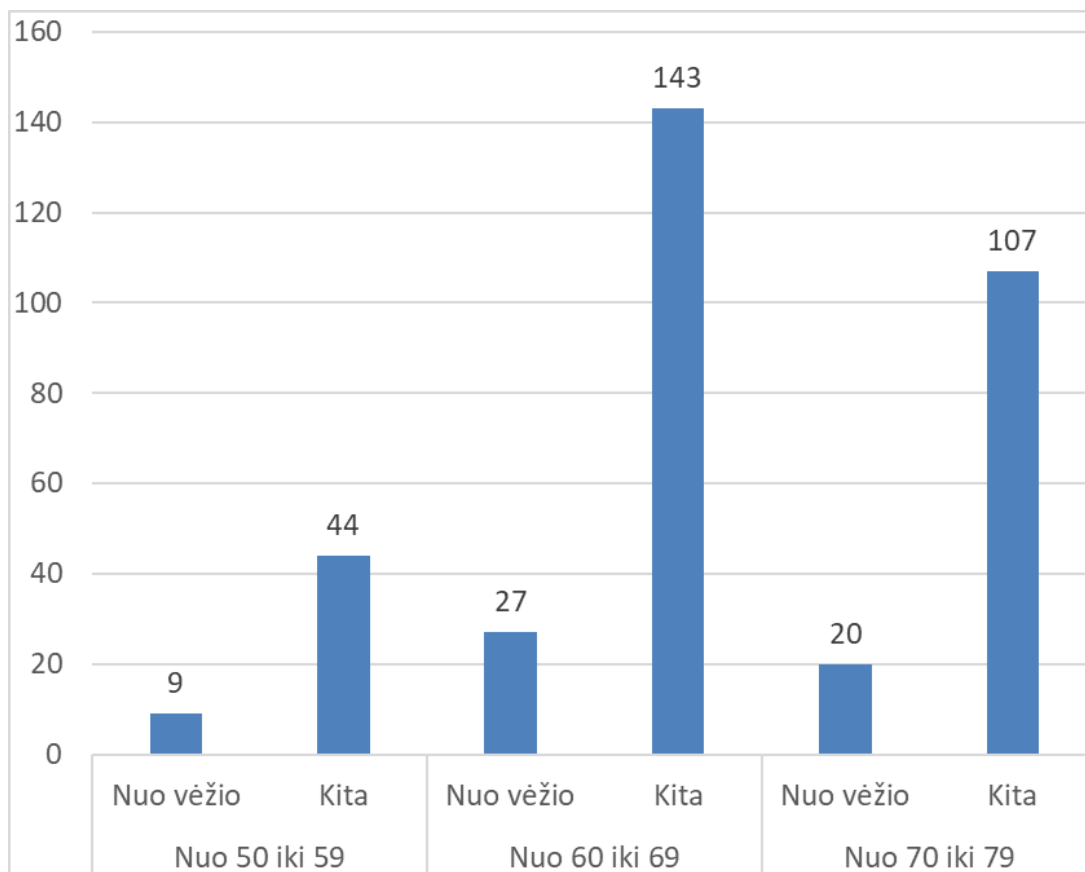
R1      SURVIVAL      OS      CSS      Death_other_causes Age64vs65
0:1606  Min.   : 1.00    0:2060  0:2354  0:2116                0:1306
1: 698  1st Qu.: 50.00    1: 350  1:  56  1: 294                1:1104
2: 106  Median : 85.00
        Mean   : 89.81
        3rd Qu.:125.00
        Max.   :206.00

```

4 priedas Skirtingų amžiaus grupių pacientų mirties rodiklis, vienetais



5 priedas Neišgyvenusių skirtingų amžiaus grupių pacientų mirties priežastis, vienetais



6 priedas Internetinio puslapio „PLV“ skirtuko vaizdas

PLV

PLV Naudojimosi instrukcija Prognozavimo įrankis Kontaktai

Priešinės liaukos vėžys

Pagal nacionalinio vėžio institutą šiuo metu Lietuvoje priešinės liaukos vėžys yra dažniausia vyrų onkologinė liga. Ši liga vyrams per metus nustatoma netgi apie 2000 naujų atvejų, o sergamumas šia liga nuolat auga. Priešinės liaukos vėžys dažniausiai aptinkamas vyrams esantiems virš 50 metų amžiaus, tuo tarpu jaunesniems aptinkama rečiau. Būtent ši liga yra labai klaidinga tuo atžvilgiu, nes ji, kitaip negu kitų organų vėžys, gali ilgus metus tūnoti priešinėje liaukoje nesukeldamas jokių simptomų. Dėl dažnai pasitaikančio lėto navikų augimo jaunesnio nei 50 metų amžiaus vyrai ši liga gali anksti nesukelti problemų. Tačiau yra atvejų, kai būtent šis vėžys vystosi sparčiai ir kai kuriais atvejais gali išplisti ir į kitus organus, ypač kaulus.

Prostata yra liauka, randama tik vyrų organizmuose. Prostata yra žemiau šlapimo pūslės (tuščiaavidurio organo, kuriame laikomas šlapimas) ir prieš tiesiąją žarną (paskutinė žarnyno dalis). Už prostatos yra liaukos, vadinamos sėklinėmis pūslėmis, iš kurių susidaro didžioji spermos skystio dalis. Šlaplė, kuri yra vamzdelis, per šlapimą ir spermą iš organizmo išvedantis per varpą, eina per prostatos centrą. Prostatos vėžys prasideda, kai prostatos liaukos ląstelės pradeda nekontroliuojamai augti. Vyrui senstant prostatos dydis gali pasikeisti. Jaunesniems vyrams jis yra maždaug graikinio riešuto dydžio, tuo tarpu vyresniems vyrams jis gali būti daug didesnis.

Beveik visi prostatos vėžiai yra adenokarcinomos. Šis vėžio tipas išsivysto iš liaukos ląstelių, tai yra ląstelių, iš kurių susidaro prostatos skystis, pridedamas prie spermos. Kiti, rečiau pasitaikantys, vėžio tipai, kurie gali prasidėti prostatoje, yra šie:

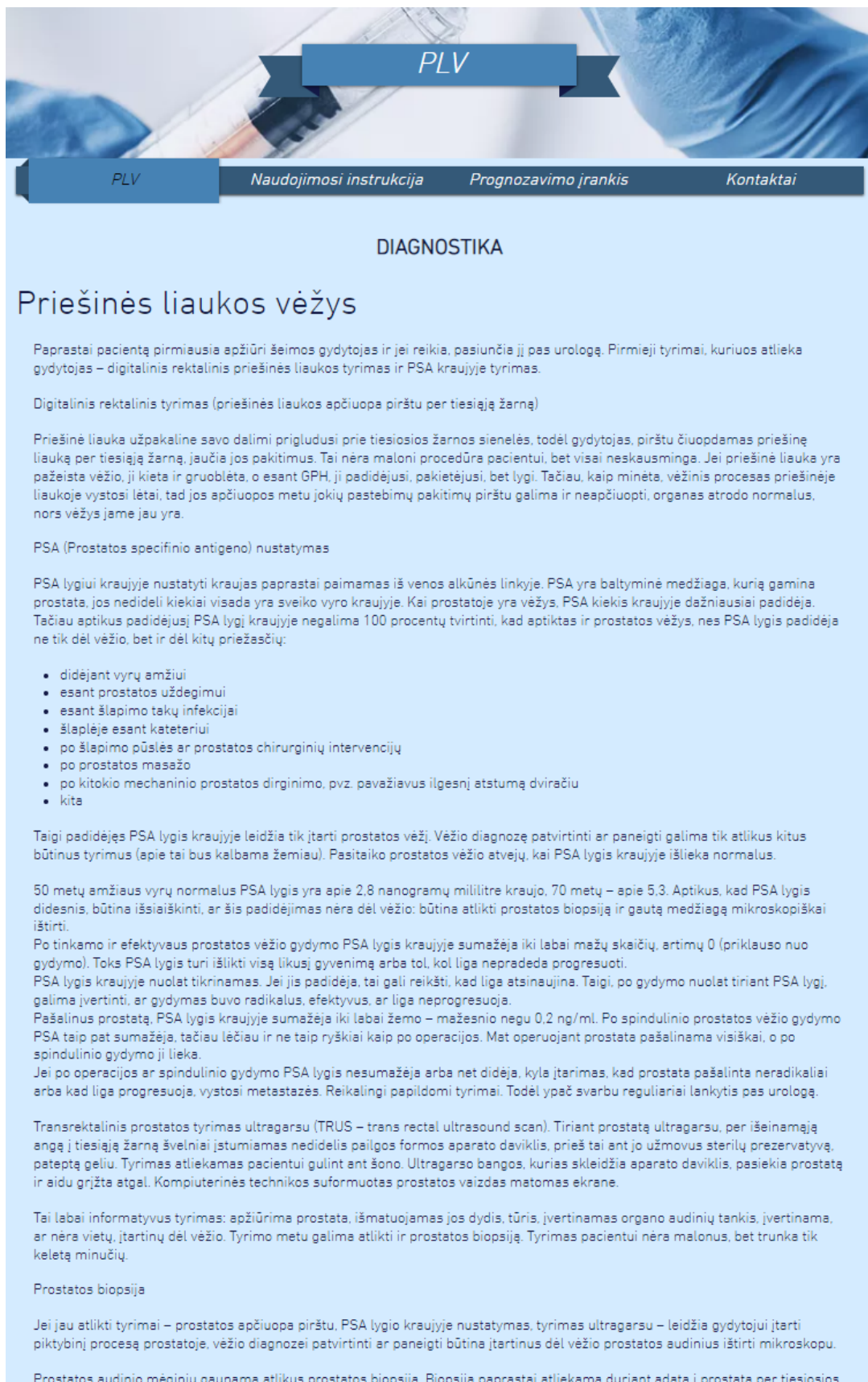
- Smulkiųjų ląstelių karcinomos
- Neuroendokrininiai navikai (išskyrus smulkiųjų ląstelių karcinomas)
- Pereinamųjų ląstelių arba urotelinis
- Sarkoma

Daugiau informacijos: <https://www.nvi.lt>

© 2021 Indrė Juškaitė. Priešinės liaukos vėžio rizikos vertinimo programa

Šiame lūžiančiame etape! 2014-2020 metų Europos Sąjungos struktūrinės paramos programos

7 priedas priedas Internetinio puslapio „Diagnostika“ skirtuko vaizdas 1



PLV

PLV **Naudojimosi instrukcija** **Prognozavimo įrankis** **Kontaktai**

DIAGNOSTIKA

Priešinės liaukos vėžys

Paprastai pacientą pirmiausia apžiūri šeimos gydytojas ir jei reikia, pasiunčia jį pas urologą. Pirmieji tyrimai, kuriuos atlieka gydytojas – digitalinis rektalinis priešinės liaukos tyrimas ir PSA kraujyje tyrimas.

Digitalinis rektalinis tyrimas (priešinės liaukos apčiuopa pirštu per tiesiąją žarną)

Priešinė liauka užpakaline savo dalimi priglundusi prie tiesiosios žarnos sienelės, todėl gydytojas, pirštu čiupuodamas priešinę liauką per tiesiąją žarną, jaučia jos pakitimus. Tai nėra maloni procedūra pacientui, bet visai neskausminga. Jei priešinė liauka yra pažeista vėžio, ji kieta ir gruoblėta, o esant GPH, ji padidėjusi, pakietėjusi, bet lygi. Tačiau, kaip minėta, vėžinis procesas priešinėje liaukoje vystosi lėtai, tad jos apčiuopos metu jokių pastebimų pakitimų pirštu galima ir neapčiuopti, organas atrodo normalus, nors vėžys jame jau yra.

PSA (Prostatos specifinio antigeno) nustatymas

PSA lygiui kraujyje nustatyti kraujas paprastai paimamas iš venos alkūnės linkyje. PSA yra baltyminė medžiaga, kurią gamina prostata, jos nedideli kiekiai visada yra sveiko vyro kraujyje. Kai prostatoje yra vėžys, PSA kiekis kraujyje dažniausiai padidėja. Tačiau aptikus padidėjusį PSA lygį kraujyje negalima 100 procentų tvirtinti, kad aptiktas ir prostatos vėžys, nes PSA lygis padidėja ne tik dėl vėžio, bet ir dėl kitų priežasčių:

- didėjant vyrų amžiui
- esant prostatos uždegimui
- esant šlapimo takų infekcijai
- šlaplėje esant kateteriui
- po šlapimo pūslės ar prostatos chirurginių intervencijų
- po prostatos masažo
- po kitokio mechaninio prostatos dirginimo, pvz. pavažiavus ilgesnį atstumą dviračiu
- kita

Taiigi padidėjęs PSA lygis kraujyje leidžia tik įtarti prostatos vėžį. Vėžio diagnozę patvirtinti ar paneigti galima tik atlikus kitus būtinus tyrimus (apie tai bus kalbama žemiau). Pasitaiko prostatos vėžio atvejų, kai PSA lygis kraujyje išlieka normalus.

50 metų amžiaus vyrų normalus PSA lygis yra apie 2,8 nanogramų mililitre kraujo, 70 metų – apie 5,3. Aptikus, kad PSA lygis didesnis, būtina išsiaiškinti, ar šis padidėjimas nėra dėl vėžio: būtina atlikti prostatos biopsiją ir gautą medžiagą mikroskopiškai ištirti.

Po tinkamo ir efektyvaus prostatos vėžio gydymo PSA lygis kraujyje sumažėja iki labai mažų skaičių, artimų 0 (priklauso nuo gydymo). Toks PSA lygis turi išlikti visą likusį gyvenimą arba tol, kol liga nepradedą progresuoti.

PSA lygis kraujyje nuolat tikrinamas. Jei jis padidėja, tai gali reikšti, kad liga atsinaujina. Taiigi, po gydymo nuolat tiriant PSA lygį, galima įvertinti, ar gydymas buvo radikalus, efektyvus, ar liga neprogresuoja.

Pašalinus prostatą, PSA lygis kraujyje sumažėja iki labai žemo – mažesnio negu 0,2 ng/ml. Po spindulinio prostatos vėžio gydymo PSA taip pat sumažėja, tačiau lėčiau ir ne taip ryškiai kaip po operacijos. Mat operuojant prostatą pašalinama visiškai, o po spindulinio gydymo ji lieka.

Jei po operacijos ar spindulinio gydymo PSA lygis nesumažėja arba net didėja, kyla įtarimas, kad prostatą pašalinata neradikalčiai arba kad liga progresuoja, vystosi metastazės. Reikalingi papildomi tyrimai. Todėl ypač svarbu reguliariai lankytis pas urologą.

Transrektalinis prostatos tyrimas ultragarsu (TRUS – trans rectal ultrasound scan). Tiriant prostatą ultragarsu, per iššėtinąją angą į tiesiąją žarną švelniai įstumiama nedidelis pailgos formos aparato daviškis, prieš tai ant jo užmovus sterilų prezervatyvą, pateptą geliu. Tyrimas atliekamas pacientui gulint ant šono. Ultragarso bangos, kurias sklaidžia aparato daviškis, pasiekia prostatą ir aidu grįžta atgal. Kompiuterinės technikos suformuotas prostatos vaizdas matomas ekrane.

Tai labai informatyvus tyrimas: apžiūrima prostata, išmatuojamas jos dydis, tūris, įvertinamas organo audinių tankis, įvertinama, ar nėra vietų, įtartinų dėl vėžio. Tyrimo metu galima atlikti ir prostatos biopsiją. Tyrimas pacientui nėra malonus, bet trunka tik keletą minučių.

Prostatos biopsija

Jei jau atlikti tyrimai – prostatos apčiuopa pirštu, PSA lygio kraujyje nustatymas, tyrimas ultragarsu – leidžia gydytojui įtarti piktybinį procesą prostatoje, vėžio diagnozei patvirtinti ar paneigti būtina įtartinus dėl vėžio prostatos audinius ištirti mikroskopu.

Prostatos audinio mėniniu gaunama atlikus prostatos biopsiją. Biopsija paprastai atliekama duriant adata į prostatą per tiesiosios

8 priedas Internetinio puslapio „Diagnostika“ skirtuko vaizdas 2

isūti.

Po tinkamo ir efektyvaus prostatos vėžio gydymo PSA lygis kraujyje sumažėja iki labai mažų skaičių, artimų 0 (priklauso nuo gydymo). Toks PSA lygis turi išlikti visą likusį gyvenimą arba tol, kol liga nepradedą progresuoti.

PSA lygis kraujyje nuolat tikrinamas. Jei jis padidėja, tai gali reikšti, kad liga atsinaujina. Taigi, po gydymo nuolat tiriant PSA lygį, galima įvertinti, ar gydymas buvo radikalus, efektyvus, ar liga neprogresuoja.

Pašalinus prostatą, PSA lygis kraujyje sumažėja iki labai žemo – mažesnio negu 0,2 ng/ml. Po spindulinio prostatos vėžio gydymo PSA taip pat sumažėja, tačiau lėčiau ir ne taip ryškiai kaip po operacijos. Mat operuojant prostatą pašalinama visiškai, o po spindulinio gydymo ji lieka.

Jei po operacijos ar spindulinio gydymo PSA lygis nesumažėja arba net didėja, kyla įtarimas, kad prostata pašalinta neradikaliai arba kad liga progresuoja, vystosi metastazės. Reikalingi papildomi tyrimai. Todėl ypač svarbu reguliariai lankytis pas urologą.

Transrektalinis prostatos tyrimas ultragarsu (TRUS – trans rectal ultrasound scan). Tiriant prostatą ultragarsu, per išeinamąją angą į tiesiąją žarną švelniai įstumiama nedidelis pailgos formos aparato daviklis, prieš tai ant jo užmovus sterilų prezervatyvą, patęptą geliu. Tyrimas atliekamas pacientui gulint ant šono. Ultragarso bangos, kurias skleidžia aparato daviklis, pasiekia prostatą ir aidu grįžta atgal. Kompiuterinės technikos suformuotas prostatos vaizdas matomas ekrane.

Tai labai informatyvus tyrimas: apžiūrima prostata, išmatuojamas jos dydis, tūris, įvertinamas organo audinių tankis, įvertinama, ar nėra vietų, įtartinų dėl vėžio. Tyrimo metu galima atlikti ir prostatos biopsiją. Tyrimas pacientui nėra malonus, bet trunka tik keletą minučių.

Prostatos biopsija

Jei jau atlikti tyrimai – prostatos apčiuopa pirštu, PSA lygio kraujyje nustatymas, tyrimas ultragarsu – leidžia gydytojui įtarti piktybinį procesą prostatoje, vėžio diagnozei patvirtinti ar paneigti būtina įtartinus dėl vėžio prostatos audinius ištirti mikroskopu.

Prostatos audinio mėginių gaunama atlikus prostatos biopsiją. Biopsija paprastai atliekama duriant adata į prostatą per tiesiosios žarnos sienelę, veiksmus kontroliuojant ultragarsu, t. y. prostatos tyrimo ultragarsu metu. Visa, kas atliekama, matoma ultragarso aparato ekrane.

Į prostatą duriama specialia adata, [dėta į vadinamąjį "pistoletą", "iššovus" adatą, ji įsminga į numatytą prostatos vietą, automatiškai paimamas audinio mėginys. Adata ištraukiama, mėginys paskleidžiamas ant objekcinio stiklelio. Tokiu būdu atliekama 6–10, kartais ir daugiau "šūvių", kaskart paimamas mėginys vis iš kitos prostatos vietos. Visus mėginius iš įvairių prostatos vietų patologas ištiria mikroskopu, išvadas pateikia po kelių dienų.

Visa biopsijos procedūra atliekama vietinės nejautos sąlygomis, į tiesiosios žarnos spindį įšvirkštus vietinio veikimo anestetikų. Po biopsijos infekcijos rizikai sumažinti pacientui skiriami antibiotikai. Parą ar dvi po biopsijos gali būti kraujo šlapinantis ar tuštinantis.

Kiekvienu atveju prieš atliekant prostatos biopsiją, gydytojas pacientui paaiškina, kaip reiktų pasiruošti procedūrai, kaip ji bus atliekama, kokių gali būti komplikacijų ir kaip jų išvengti. Gaila, bet 20 % atvejų prostatos vėžio nepavyksta įrodyti ir atlikus biopsiją bei ištyrus medžiagą mikroskopu. Tokiais atvejais nors biopsijos rezultatai ir negatyvūs, įtarimas dėl vėžio lieka, todėl praėjus tam tikram laikui biopsiją reikia pakartoti.

Kiti tyrimai

Jei nustatytas prostatos vėžys, būtina įvertinti bendrą paciento sveikatos būklę bei įsitikinti, ar liga neišplito už organo ribų – reikia atlikti kitus tyrimus. Nebūtinai visi toliau minimi tyrimai atliekami kiekvienu prostatos vėžio atveju. Kokie tyrimai reikalingi, sprendžia gydytojas.


- Rentgeniniai tyrimai
- Intraveninė urograma
- Kaulų skenavimas. Kaip minėta, prostatos vėžys greičiausiai metastazuoja į kaulus. Šis tyrimas atliekamas nustačius prostatos vėžį, kai yra didelė metastazių kauluose tikimybė – aukštas PSA lygis kraujyje, pacientas skundžiasi kaulų skausmais, gydytojas pirštu apčiuopia didelę, grublėtą prostatą ir kt.
- Pilvo organų ultragarsinis tyrimas
- Kompiuterinė tomografija (KT). Prostatos vėžio atveju gali būti atlikta dubens srities organų KT, tačiau kai reikia, atliekama ir kitų sričių KT. Kai atliekama dubens srities organų KT, pacientui keletą valandų prieš tyrimą duodama išgerti specialios kontrastinės medžiagos, kad KT nuotraukų vaizdai būtų aiškesni ir lengviau įvertinami.
- Magnetinio rezonanso tomografija
- Cistoskopija



© 2021 Indrė Juškaitė. Priešinės liaukos vėžio rizikos vertinimo programa




9 priedas Naudojimosi instrukcijos skirtuko vaizdas 1



PLV

PLV Naudojimosi instrukcija Prognozavimo įrankis Kontaktai

Kaip naudotis programa?




R Studio

1. R studio

- 1.1 Parsisiųskite R Studio programinį paketą. (nemokama versija)
Nuoroda:
<https://www.rstudio.com/products/rstudio/download/>
- 1.2 Instaliuokite programą laikydamiesi nurodymų.

Papildoma informacija apie produktą:
<https://www.r-project.org/>



PLV Naudojimosi instrukcija Prognozavimo įrankis Kontaktai


Priešinės liaukos vėžio modelio
PROGRAMA

Prieš naudojant, prašome perskaityti "Naudojimosi instrukcija".
Paspaudus mygtuką "Pradėti", Jūs būsite nukelti į puslapį kuriame yra patalpinta
programa.

Pradėti

2. Parsisiųskite "Prognozavimo įrankio" failus

- 2.1 Nueikite į "Prognozavimo įrankis" skirtuką.
- 2.2 Paspauskite mygtuką "Pradėti"



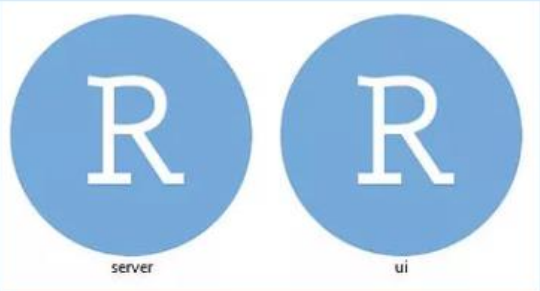
Sign in Save to Dropbox Download

PLV

Sorted by name

DATABASE server ui

- 2.3 Parsisiųskite ir išsaugokite visus failus esančius naujai atsidariusiame lange.

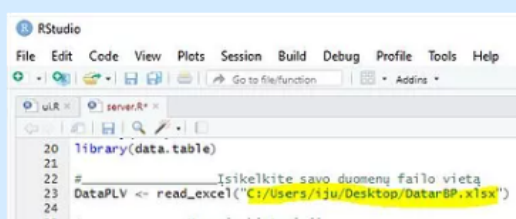


server ui

3. Atsidarykite programą

- 3.1 Du kartus paspauskite ant failo "ui".
- 3.2 Du kartus paspauskite ant failo "server"

10 priedas Naudojimosi instrukcijos skirtuko vaizdas 2

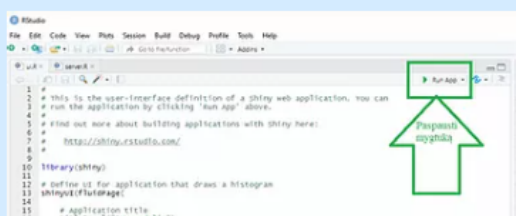


```
20 library(data.table)
21
22 # Isikelkite savo duomenų failo vietą
23 DataPLV <- read_excel("C:/Users/tju/Desktop/DatarBP.xlsx")
24
25 #
```

4. Atlaujinkite duomenų failo laikymo vietą

4.1 22 ir 23 "server.R" programos eilutėse nurodykite kur išsaugojote excel failo kelią.

4.2 Įklijuokite nukopijuoto failo kelią tarp kabučių, kaip parodyta pavyzdyje.

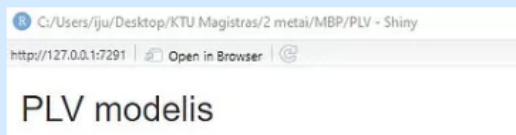


```
1 #
2 # this is the user-interface definition of a shiny web application, you can
3 # run the application by clicking "run app" above.
4 #
5 # find out more about building applications with shiny here:
6 # https://shiny.rstudio.com/
7 #
8 #
9
10 library(shiny)
11 # define UI for application that draws a histogram
12 shinyUI(fluidPage()
13   # application title
14   titlePanel("PLV modelis"),
15 )
```

5. Paleiskite įrankį

5.1 Paspauskite "Run App" mygtuką

5.2 Dirbkite atsidariusiame lange arba paspauskite "Open in Browser" ir dirbkite internetinio puslapio lange.



PLV modelis

Paciento amžius:

PSA:

Išvėjęs(-i) limfmazgiu skaičius:

Patologinė vezio stadija:

Limfmazgiu bukie:

Pasalinęs prostatos bukie:

Paspauskite mygtuką kuris pritaikys modelį.

6. Užpildykite visus laukelius

6.1 Užpildykite visus įvesties laukus.

6.2 Paspauskite mygtuką "Pritaikyti modelį".

Pastaba! Pirmą kartą pritaikius modelį teks palaukti kelias minutes, kol neuroniniai tinklai pritaikys modelius.

Rezultatai ir jų interpretavimas



PLV modelis

Paciento amžius:

PSA:

Išvėjęs(-i) limfmazgiu skaičius:

Patologinė vezio stadija:

Limfmazgiu bukie:

Pasalinęs prostatos bukie:

Paspauskite mygtuką kuris pritaikys modelį.

Su šiais rodikliais įgyvendinamas modelis.
Su šiais rodikliais įgyvendinamas modelis.
Dėkui už suinteresuotumą!

Rezultatai

Dešniame programos lango dalyje išvedami rezultatai. Virš kiekvieno rezultato išspausdintas trumpas paaiškinimas.

11 priedas Statistinės ir koreliacinės analizės programos kodai

MBP.r

```
library(readxl)

DataMBP <- read_excel("C:/Users/iju/Desktop/DatarBP.xlsx")

#View(DataMBP)

#_____Bendroji duomenu analizes dalis_____

str(DataMBP)

DataMBP$GLpoRP_ISUP <- as.factor(DataMBP$GLpoRP_ISUP)

DataMBP$pT <- as.factor(DataMBP$pT)

DataMBP$LNI01 <- as.factor(DataMBP$LNI01)

DataMBP$R1 <- as.factor(DataMBP$R1)

DataMBP$OS <- as.factor(DataMBP$OS)

DataMBP$CSS <- as.factor(DataMBP$CSS)

DataMBP$Death_other_causes <- as.factor(DataMBP$Death_other_causes)

DataMBP$Age64vs65 <- as.factor(DataMBP$Age64vs65)

str(DataMBP)

summary(DataMBP)

#_____ Numeric koreliacijai _____

DataMBP$Nr <- as.numeric(as.factor(DataMBP$Nr))

DataMBP$Amzius <- as.numeric(as.factor(DataMBP$Amzius))

DataMBP$PSAI <- as.numeric(as.factor(DataMBP$PSAI))

DataMBP$GLpoRP_ISUP <- as.numeric(as.factor(DataMBP$GLpoRP_ISUP))

DataMBP$pT <- as.numeric(as.factor(DataMBP$pT))

DataMBP$LNI01 <- as.numeric(as.factor(DataMBP$LNI01))

DataMBP$R1 <- as.numeric(as.factor(DataMBP$R1))

DataMBP$SURVIVAL <- as.numeric(as.factor(DataMBP$SURVIVAL))

DataMBP$OS <- as.numeric(as.factor(DataMBP$OS))

DataMBP$CSS <- as.numeric(as.factor(DataMBP$CSS))

DataMBP$Death_other_causes <- as.numeric(as.factor(DataMBP$Death_other_causes))

DataMBP$Age64vs65 <- as.numeric(as.factor(DataMBP$Age64vs65))

#_____ normal distribution, tikrinimas

ggdensity(DataMBP$Amzius)
```

```

ggdensity(DataMBP$PSAI)
ggdensity(DataMBP$GLpORP_ISUP)
ggdensity(DataMBP$pT)
ggdensity(DataMBP$LNI01)
ggdensity(DataMBP$R1)
ggdensity(DataMBP$SURVIVAL)
ggdensity(DataMBP$OS)
ggdensity(DataMBP$CSS)
ggdensity(DataMBP$Death_other_causes)
ggdensity(DataMBP$Age64vs65)

hist(DataMBP$Amzius, main = "Amžiaus pasiskirstymo grafikas",
      xlab="Amžius",
      ylab="Dažnis"
      )

hist(DataMBP$PSAI, main = "PSA pasiskirstymo grafikas",
      xlab="PSA",
      ylab="Dažnis",
      xlim=c(0,60),
      ylim=c(0,2000))

hist(DataMBP$SURVIVAL, main = "Išgyvenamumo pasiskirstymo grafikas",
      xlab="Išgyvenamumas",
      ylab="Dažnis",
      xlim=c(0,200),
      ylim=c(0,400))

#plot(DataMBP$PSAI ~ DataMBP$SURVIVAL, data = DataMBP)

#_____koreliacija pirsono
corData <- DataMBP[,]
corPirson <- cor(corData)

library(corrplot)
library(corrgram)

corrplot(corrgram(corData))

write.xlsx(corPirson, file = "DATAKORELIACIJAAS.xlsx", sheetName = "Sheet1",

```

```

        col.names = TRUE, row.names = TRUE, append = FALSE)
write.xlsx(corPirson, file = "DATAKORELIACIJAAS.xlsx")
#_____koreliacija pirsono
corData <- DataMBP[c(-1)]
corPirson <- cor(corData)
corrplot(corrgram(corPirson))
#_____koreliacija kendal diskretiems kint
corKendal <- cor(corData[c(3, 4, 5, 6, 8, 9, 10, 11)], method="kendall")
corrplot(corrgram(corKendal))
#_____koreliacija spearman tolygiems kint.
corSpearman <- cor(corData[c(1, 2, 7)], method="spearman")
corrplot(corrgram(corSpearman))
# Using cor.test() method
res <- cor.test(DataMBP$Amzius, DataMBP$PSAI, method = "spearman")
print(res)

```